

# THIS WEEK



## EDITORIALS

**EGYPT** Opening a fresh chapter for the Bibliotheca Alexandria **p.6**

**WORLD VIEW** A newspaper reporter's take on mutant flu research **p.7**

**TRUE BLUE** Synthetic material wings to perfect colour **p.8**

## Publishing risky research

*Imperfect global biosafety standards and a threat to researchers' motivations from biosecurity concerns are among the significant risks in current flu research.*

This week sees the online publication of the paper 'Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets' by the Japanese-US team headed by Yoshihiro Kawaoka at the University of Wisconsin-Madison (M. Imai *et al.* *Nature* <http://dx.doi.org/10.1038/nature10831> (2012). See also pages 7 and 13, and H.-L. Yen and J. S. M. Peiris *Nature* <http://dx.doi.org/10.1038/nature11192>; 2012). Kawaoka's paper was one of two submitted last August, reporting mammalian transmissibility of avian flu as a result of artificial genetic manipulation, the principal scientific interest of which arises from the small number of mutations found to be necessary. The other paper, by a team headed by Ron Fouchier at the Erasmus Medical Centre in Rotterdam, the Netherlands, is expected to appear soon in *Science*.

As has been much discussed in *Nature*, both papers were independently assessed by the US National Science Advisory Board for Biosecurity (NSABB) while being considered by the journals. The NSABB's recommendation, communicated to the journals in November last year, was not to publish the essential methods and data. Although such a recommendation has no statutory force, it makes any researcher or publisher pause. There followed months of public debate and two two-day meetings involving flu experts and other stakeholders, one held by the World Health Organization (WHO) and the other by the NSABB. After the second, at the end of March, the NSABB essentially reversed its position, and *Nature* made its own decision to proceed.

### LESSONS LEARNED

As the economist John Maynard Keynes reportedly said: "When the facts change, I change my mind." But the essential scientific elements in the Kawaoka paper were unchanged between the first and second NSABB deliberations. It is now clear that the committee's original deliberations were too limited, especially given the enormous implications for flu research of a recommendation against publication. Yet as a body that aims to anticipate and scrutinize the security risks of biological research, the NSABB is unique worldwide, and it is desirable to have such a forum. The discussion that followed the board's first decision would not have been as valuable or as prompt had it concerned hypothetical cases. Yet there are justified concerns among the research community about the NSABB's processes, and these processes should be reviewed.

Some lessons have emerged that point to actions and policies for the future. First, it was worth deliberating at length on the possibility of redacting the key findings of the paper instead of simply rejecting it. (Rejection has long been an option if *Nature* is advised by security experts that the risks of publication exceed the benefits.) There was also the option that the full paper might be distributed by some third party, to selected recipients only. Having now considered these matters in depth, the editors of this journal have decided that we will not consider either alternative for papers in *Nature* in the foreseeable future. A paper that omits key results or methods disables subsequent

research and peer review. Furthermore, after much internal and external deliberation, we cannot imagine any mechanism or criterion by which to sensibly judge who should or should not be allowed to see the work. Nor do we believe that any restricted information distributed to university laboratories would stay confidential for long.

We are aware that the lack of an option for restricted publication has its own risks in a discipline in which results might be both beneficial to the public benefit and a threat to security. We will willingly explore ways out of this dilemma.

**"Any restricted information distributed to university laboratories would not stay confidential for long."**

One major risk amid these discussions is that younger researchers might be discouraged from entering a field that is subject to security constraints. But the attitudes of biosecurity experts are more encouraging than is widely appreciated.

As far as *Nature* is aware, formal assessments by security agencies have led to recommendations that the Kawaoka paper be published.

This includes an independent assessment that we commissioned from a non-US biological-defence agency, whose advice can be found at [go.nature.com/wglsea](http://go.nature.com/wglsea). In subsequent discussions with biosecurity researchers, there has been a striking unanimity: where there is a benefit to public health or science, publish! It has been enlightening to see how scientists in this secretive arena see the open scientific enterprise as their best recourse in times of potential trouble.

The third most important lesson is about biosafety. Here there are real concerns: humans lack immunity to flu viruses with an H5 haemagglutinin protein, and an accidental release of a mammalian-transmissible H5 virus would have the potential to cause a pandemic were it to transmit between humans. A key component of the second round of NSABB deliberations was a clear presentation by Kawaoka of his team's very rigorous security processes and set-up, including physical arrangements, training and due diligence exercised with personnel.

Such a reassuring picture is not globally applicable. The standards of these labs (fully described in the Kawaoka paper) were widely quoted as biosafety level (BSL) 3 enhanced. The WHO discussion considered such standards essential, and worried that to require the distinctly more demanding BSL-4 standard would shut down the research. However, 'BSL-3 enhanced' is not a formally established standard. What is more, not every country may have sufficient regulatory systems and robust laboratory cultures of safety. This is a key issue as the self-imposed moratorium on work by flu-transmissibility researchers continues.

The WHO will soon release guidelines about international standards for biosafety. The signs are that these will highlight key issues and aspects of good governance, but will not themselves provide a framework for strengthened implementation. The absence of such a framework is an urgent concern for all researchers working with dangerous organisms, and for all who fund and publish their work. ■

# Quiet please

*Ismail Serageldin deserves the chance to prepare a new future for the Alexandria library.*

One of the defining images of the revolution in Egypt last year was of people linking hands around the Bibliotheca Alexandrina, the modern incarnation of the renowned Library of Alexandria destroyed in antiquity. At the height of the unrest, young people from the city formed a 24-hour protective cordon to save the library from being over-run by angry crowds.

The Bibliotheca, built close to the site of the original library, is popular with Egyptians, around a million of whom visit the library, the manuscript archives, and the associated museums and galleries each year. The staff organize hundreds of science clubs in Alexandrian schools, and scholars from all over the world give talks and workshops. Its lecture halls are packed most nights and the BioVision life-sciences conference held there last month was massively oversubscribed.

Alexandrians are immensely proud of their library, but some are sceptical of its director, Ismail Serageldin. A long-time champion of rationalism and reform in the Arab world, Serageldin should be an influential voice in the new Egypt. And yet when he speaks, few listen. What offends people most is his closeness to the hated Mubarak.

Young and old in post-revolution Egypt have found the voice denied to them in 30 years of dictatorial rule and a common mantra is of fresh starts. They will not tolerate the continued presence of people and institutions that prospered during a regime that tortured its own people. Mubarak-era loyalists who seek a place in the new Egypt are parodied as *fulool* — turncoats. Serageldin is no *fulool*, but his record as a voice for secular reform is not of much consequence now. For

his critics, what matters is that Mubarak's wife Suzanne chaired the library's board of trustees and that Serageldin accepted a seat in the Egyptian Senate at the request of her husband.

That is why there have been calls in the newly elected parliament for Serageldin to stand down. And it is why, on most afternoons, a small group of protesters stands outside the Bibliotheca, holding a single poster. In red letters are the words "Please leave it". Below the text is an enlarged photograph of Serageldin, his face crossed-out with a large X.

Serageldin retains the support of the library's 26-strong board of international trustees. In practice, however, a decision on his future will be made by Islamist MPs, who dominate the parliament, and by whoever wins the presidency in June. But they should remember that Serageldin's predicament has many parallels in Islamic history.

**"Serageldin's predicament has many parallels in Islamic history."**

During the golden age of Islamic science, scientific advances were often associated with repression, principally because scholars had to rely on the reigning despots to support and fund their work. Important contributions to algebra and optics, for example, were made during regimes in the ninth and eleventh centuries AD that were repressive even by the standards of the time. And yet the scientists who worked for them, such as the mathematician al-Khwarizmi in Baghdad and the polymath ibn al-Haytham in Cairo, are now celebrated as pathfinders. Similarly, Serageldin could not have achieved what he did at the library without the support of the Mubarak regime.

An elected government would be well within its rights to start a new chapter in the library's leadership now. But a wiser course of action might be to allow Serageldin to complete his current term of office and to prepare a succession plan for the end of his directorship in 2015. That would take a lesson from history and honour the contributions of a talented and innovative leader. ■

# Young Americans

*The rancorous debate over when people first arrived in America has not helped science.*

In 1781, American statesman Thomas Jefferson noted that Asia and America are separated "only by a narrow strait" that would have allowed passage between the two continents. There is no record of any scholarly scorn of Jefferson's ideas about the peopling of America. But in recent decades, anyone wading into the topic has needed skin as thick as a woolly mammoth's. The debate over the first Americans has been one of the most acrimonious — and unfruitful — in all of science.

As reported on page 30, one side of that debate held that, around 13,000 years ago, a group of ice-age hunters from Siberia crossed into North America over Jefferson's "narrow strait" — which, at the time, was an exposed strip of land. Armed with stone weapons, called Clovis points, these hunters spread rapidly across the continent and feasted on animals that had never known humans. The opposing camp argued that people reached the Americas long before Clovis technology appeared — at least 1,000 years earlier.

The histories of these arguments are a case study of poor communication and missed opportunities. One researcher, new to the field after years of working on other contentious topics, told *Nature* that he had never before witnessed the level of aggression that swirled around the issue of who reached America first. "When people stop listening to arguments and stop looking at data and instead just go with their own beliefs," he said, "that's when it becomes completely crazy."

He was referring to researchers who support the Clovis-first model,

which was the dominant hypothesis from the second half of the twentieth century until only a few years ago. Researchers who went against that model by reporting even older sites of human occupation endured brutal criticism from opponents who did not give them, or their evidence, a fair hearing. Scientists who supported the Clovis-first model countered that reports of pre-Clovis sites were examples of poor scholarship.

Studies from the past few years now offer a convincing case that humans reached the Americas well before the Clovis culture. Credit for this breakthrough should go to open-minded archaeologists, who were willing to investigate pre-Clovis sites seriously, and to geneticists, for bringing fresh ideas and techniques to bear on the topic. The recent finds and the shift in the debate have triggered a renaissance in ancient-American archaeology. Researchers are reopening sites, re-examining specimens and searching for new sites to determine who the early pioneers were, and how and when they arrived.

As these ancient events are explored, some archaeologists should examine their recent behaviour. If what they lacked could be summed up in one word, it might be respect. Researchers must always consider that they might be wrong, and should look carefully at opposing data and conclusions. At the same time, scientists who make bold claims must marshal an extraordinary case, especially if they seek to topple a dominant model built on many previous studies. Such prescriptions sound obvious, but many scientists forget them, particularly in fields with limited data, such as archaeology.

The various factions could take lessons from those early Americans (whoever they were). These ancient adventurers spurned the well-trodden paths of their forebears to explore new territory. But they would not have succeeded without the help of those who came before and gave them their start. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunjv](http://go.nature.com/xhunjv)





## Mutant flu — the view from the newsroom

When reporters aren't given the facts, they are likely to jump to the most extreme conclusions, says David Brown.

When the federal building in Oklahoma City was bombed in 1995 by antigovernment terrorists, the newspaper I write for, *The Washington Post*, ran a story about how easy it was to make a bomb from nitrogenous fertilizer and motor oil. The paper's editors had debated over whether they should run the story, and whether it should be put on the website — which was very new — thereby rendering it easily accessible forever. They decided to put it up.

A lot of dangerous information is now within a few key strokes of the reading public. We hear that all the information needed to make a nuclear weapon is on the Internet (although I am not competent to judge). The base sequences of the Spanish flu, smallpox and polio viruses are publicly available. So how should we in the press view the debate over whether or not *Nature* and *Science* should have been free to publish research on mutant flu viruses? Covering it as a news story has raised many interesting issues, and that is unlikely to change with the publication of one of the papers by *Nature* (M. Imai *et al.* *Nature* <http://dx.doi.org/10.1038/nature10831>; 2012).

The concerns that reporters and newspaper editors have as they approach a story are not necessarily the same that scientists, journal editors, guardians of public health and government officials have. But there is a lot of overlap.

The first and most obvious point is that journalists' sources of knowledge are what they see, what they read and what they are told. We have no special channels. The debate over how to publish these two papers illustrates what happens when two of these sources are missing. We had no events to observe and as long as the papers remained secret we could not read them. What we were left with is what people had told us.

In my experience, scientists are more capable of giving a disinterested version of events than, say, politicians. They are willing to present both sides more or less fairly. They tend to have a respect for facts, wherever they lead, which cannot be said of lots of people who make up the voices on the news. But their general goodwill can't make up for the fact that what they told reporters about the mutant-flu controversy was woefully inadequate. The news stories that appeared were based on second- and third-hand versions of the facts. There was no evidence base for reporters to consult. And without numbers and details — the evidence base for science reporting — you get qualitative statements and metaphors. "Doomsday virus", for example.

Many reporters, it should be noted, are enthusiastic participants in overstating the case. Some learn early to make the most extreme assertions supportable by the facts even if the impression left by their story is not what one would get after a complete survey of the facts. Editors sometimes goad reporters in that direction, but generally that's not necessary.

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/pmc9nn](http://go.nature.com/pmc9nn)

It may turn out that in some circumstances, censorship is the responsible act because the facts are too dangerous to provide to the readers. But when that happens, I think you can count on getting an extreme account in the press of what is being kept secret.

Secrecy breeds paranoia. It makes people question motives. I haven't read any paranoid explanations of why this research was kept secret. But if this starts to become a habit with dual-use research, I am sure they will emerge. Conspiracy theories have a corrosive effect on democracy. They also have particular resonance in minority communities, which tend to distrust motives, and arguably science, more than society as a whole.

Secrecy also puts the press in the awkward position of delivering a message that it can't judge the credibility of, a message that says: "This information is too dangerous for all but a few experts to know — at least that's what they are telling us."

Very little is as viscerally repellant to reporters as being forced to say something like that. Nobody wants to be played for a patsy. When reporters find themselves in the position of having to write an incomplete story, they may start to question the motives and competence of the people delivering the scraps of information. (Newspapers can, however, keep secrets when they need to, such as when their own reporters are kidnapped in war zones.)

Secrecy is difficult to maintain. When the H5N1 story broke, I suggested to my editor that the newspaper might want to see if one of our computer experts could find the censored

papers in the digital cloud. We were not, of course, prepared to do any hacking to find them. But even the suggestion of searching for them made people in the newsroom very uncomfortable. "It just doesn't feel right," said one of the editors. "What are we going to do if we find them?" There was so much ambivalence that the idea just died. But if even what some people like to call the 'elite' media have compunctions about making dangerous information public, or about rooting around for it, then I don't think emerging forms will have a similar reluctance.

I am glad that the research papers are to be published in full. Attempted secrecy swims against the tide of history. For better or worse — and I think most people would say for better — paternalism has been rejected in medicine and is being challenged in other spheres, such as education policy. People expect to be able to make up their own minds, even about things they don't understand. The press should be a vehicle to help people make these decisions. ■ [SEE EDITORIAL P.6](#)

David Brown is a reporter at The Washington Post. This article is adapted from a talk originally given at the Royal Society in London. e-mail: [browndm@washpost.com](mailto:browndm@washpost.com)

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## CLIMATE MODELLING

### Wetter, drier with warming

Trends in tropical rainfall over land could provide a signal of global warming, with very wet and very dry months both projected to become more frequent as temperatures rise.

Benjamin Lintner at Rutgers University in New Brunswick, New Jersey, and his colleagues compared the results from climate models simulating rainfall for 1975–99 and 2075–99. The researchers found that dry areas will tend to become drier by the end of this century as Earth warms, whereas rainfall will increase in wet areas.

The team also compared rainfall simulations with measurements obtained between 1979 and 2008. However, there was no clear link between changes in tropical precipitation and recent temperature increases. This suggests that natural climate variability, in addition to global warming, is important in influencing tropical rainfall trends.

*J. Geophys. Res.* <http://dx.doi.org/10.1029/2012JD017499> (2012)

## BIOENGINEERING

### Shunting bacteria on a chip

A device that moves chemical reagents through a network of tiny chambers and channels on

a chip has been used to analyse the genomes of individual bacterial cells.

Carl Hansen at the University of British Columbia in Vancouver, Canada, and his team developed the microfluidic device, which uses computer-controlled valves

to pump reagents and cells as discrete droplets suspended in a stream of oil. By programming the valves, the authors directed single droplets to 95 different 30-nanolitre reaction wells (pictured, with

droplet), in which the droplets can be isolated or mixed with others. The authors used their device to sort, culture and sequence the genomes of individual bacterial cells sampled from the human mouth and from marine sediment.

*Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1106752109> (2012)

## CANCER

### A tumour's *Kras* behaviour

Growing cancer cells divert glucose and other nutrients

away from energy-producing pathways towards those that make complex molecules that can be used as building blocks. The cancer gene *Kras* promotes this diversion in pancreatic cancer.

Using a mouse model of the most common type of pancreatic cancer, Alec Kimmelman at the Dana-Farber Cancer Institute in Boston, Massachusetts, Ronald DePinho, now at the MD Anderson Cancer Center in Houston, Texas, and their colleagues show that the pancreatic tumours need mutant *Kras* to survive. The mutant protein shunts



## MATERIALS

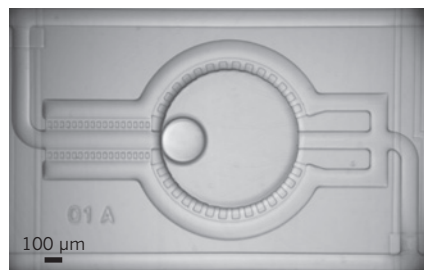
### Butterfly-inspired reflectors

A synthetic material mimics the complex microscopic structure of brilliant butterfly wings to achieve a bright blue colour.

Butterflies of the genus *Morpho* are known for their dazzling blue wings (pictured right). The colour arises from densely packed layers of ridges that cover the scales on the wing surface. The ridges are structured such that light waves reflecting off the ridges interfere with each other, creating the blue colour. The tight, semi-random packing of the ridges makes the wings appear bright across a wide range of viewing angles.

To recreate these features in a reflective material, Jung Shin at the Korea Advanced Institute of Science and Technology in Daejeon, South Korea, and his colleagues deposited silica microspheres onto a surface and then sprayed layers of titanium dioxide and silicon dioxide over them. The resulting film (pictured left) had just the right mix of regularity and disorder to create the even blue colouring.

*Adv. Mater.* <http://dx.doi.org/10.1002/adma.201200521> (2012)



KASTON LEUNG

JOHN WILEY AND SONS

glucose metabolism towards two pathways — one that adds sugars to proteins and another that makes precursors for DNA and RNA synthesis — by downregulating a glucose transporter and other key enzymes.

The enzymes altered by mutant Kras represent potential new drug targets, the authors conclude.

*Cell* 149, 656–670 (2012)

## MATERIALS

## Graphene's silicon cousin

Silicon can form two-dimensional honeycomb sheets one atom thick. This material — known as silicene — is the silicon equivalent of carbon-based graphene, and could fit more easily into industrial silicon-based circuits than graphene.

Patrick Vogt at the Technical University of Berlin, Paola De Padova at the Institute of Structure of Matter in Rome and their colleagues deposited a single layer of silicon onto a silver surface by heating a piece of silicon wafer to more than 200°C. The resulting silicene is corrugated and seems to have similar electronic properties to graphene.

Other researchers have reported making silicene before, but Vogt and colleagues say that they have provided more conclusive evidence, including microscopy images and measurements of the material's electrical and chemical properties.

*Phys. Rev. Lett.* 108, 155501 (2012)

## NEUROSCIENCE

## Small RNAs boost memory process

A class of small RNA molecules discovered just six years ago has now been detected in the brain, where it seems to regulate the expression of a gene involved in memory. This could help to explain how long-lasting

memories are maintained.

Researchers in New York led by Thomas Tuschl at the Rockefeller University and Eric Kandel at Columbia University analysed a collection of small RNAs that do not encode proteins, from the central nervous system of the sea slug *Aplysia*. The researchers identified a group of RNAs called Piwi-associated RNAs (piRNAs) that bind to a protein called Piwi. Complexes of piRNAs and Piwi protein silence a memory-inhibiting gene called CREB2 by promoting the addition of methyl groups to the gene in the presence of serotonin, a neurotransmitter important in learning and memory. Silencing CREB2 boosts the ability of neurons to change the strength of their connections in a sustained way — a key process in long-term memory.

These small RNAs had previously been found only in reproductive organs, so this study suggests a broader role for the molecules than expected.

*Cell* 149, 693–707 (2012)

## ZOOLOGY

## Jellies reproduce as little larvae

A population of comb jellies in the central Baltic Sea is the first of its kind to be discovered living and reproducing entirely in the larval stage.

*Mertensia ovum* (pictured) is common in the Arctic. Cornelia Jaspers at the Technical University of Denmark in Charlottenlund and her colleagues discovered thousands of *M. ovum* jelly larvae and hundreds of eggs during sampling cruises in the central Baltic Sea in 2009–10. They found no adults and the larvae measured at most 1.6 millimetres. However, the larvae reproduced at a rate that would sustain the observed population.

The authors suggest that this comb jelly population may have been driven to reproduce at a young age in an attempt to

## COMMUNITY CHOICE

The most viewed papers in science

## CARDIOVASCULAR BIOLOGY

## Watching risky blood clots form

**HIGHLY READ**  
on [jem.rupress.org](http://jem.rupress.org)  
24 Mar–23 Apr

Dangerous blood clots form in veins as a result of cross-talk between white blood cells and platelets. This mechanism underpins deep vein thrombosis (DVT) — a condition that can be life-threatening if the clots move to the lungs.

Steffen Massberg at the Technical University of Munich in Germany and his colleagues created a mouse model that displays many of the hallmarks of human DVT. *In vivo* imaging revealed that as blood flow is reduced — as can occur in humans during extended bouts of inactivity, such as on long-haul flights — DVT begins when monocytes and neutrophils, two types of white blood cell, stick to blood vessel walls. These cells then release molecules that trigger coagulation. Clot formation is enhanced when platelets adhere to either the vessel wall or to the attached white blood cells.

*J. Exp. Med.* 209, 819–835 (2012)

avoid the abundant predators in the region.

*Biol. Lett.* <http://dx.doi.org/10.1098/rsbl.2012.0163> (2012)

compound, norspermidine, in *Bacillus subtilis*. Jon Clardy at Harvard Medical School in Boston, Richard Losick at Harvard University in Cambridge and their colleagues found that norspermidine acts in concert with a previously identified mix of amino acids to break down established biofilms of *B. subtilis* and to prevent biofilm formation in *B. subtilis*, *Escherichia coli* and *Staphylococcus aureus*. Using high-resolution microscopy, the researchers showed that norspermidine interacts directly with the sugar-based molecules that hold the biofilm together. Mutant bacterial strains unable to make both norspermidine and the amino acids form long-lived biofilms.

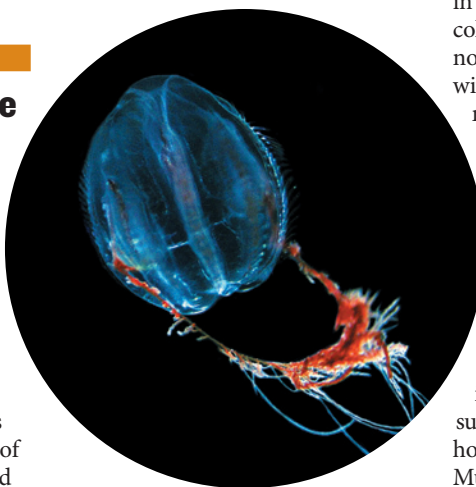
Synthetic chemicals modelled on norspermidine could be used to combat biofilms in industrial and medical settings, the authors say.

*Cell* 149, 684–692 (2012)

➔ **NATURE.COM**

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)



C. JASPERS

## MICROBIOLOGY

## Bacterial biofilm breakdown

A compound produced by certain bacteria can break down biofilms — the tight-knit communities that some bacteria form on surfaces to protect themselves from antimicrobial attack.

Researchers in Massachusetts identified the



# SEVEN DAYS

The news in brief

## EVENTS

### US finds BSE case

The United States has reported its fourth-ever case of a cow with 'mad cow disease' (bovine spongiform encephalopathy or BSE). On 24 April, the US Department of Agriculture confirmed that routine surveillance revealed the disease in a recently deceased dairy cow in central California. The infected animal carried a rare 'L-type' strain of the disease not detected before in the United States, but it never posed a threat to the human food supply, the agency said. The country's last reported BSE case was in 2006. See [go.nature.com/ntao9c](http://go.nature.com/ntao9c) for more.

## BUSINESS

### Cellulosic-fuel fail

Iogen, one of the world's leading producers of cellulosic fuels (made from straw and other plant waste), has scrapped plans to build a commercial-scale plant in southern Manitoba, Canada. The firm, which is based in Ottawa, had planned to build the plant in partnership with its joint owner, the oil giant Royal Dutch Shell. The 30 April announcement did not explain why the project was scrapped, but a municipal official in Manitoba has said that Shell told him the costs were too high.

## POLICY

### Low-carbon Japan?

Japan's government has approved a revised environmental plan that includes the eyebrow-raising target of cutting greenhouse-gas emissions by 80% by 2050 (compared with 1990 levels). The ambitious goal has been mentioned before, but got its first formal acknowledgement in the plan passed on 27 April.



YOMIURI SHIMBUN/AP

## Drilling into Japan's quake zone

A research expedition has drilled more than 7.7 kilometres below the sea surface to reach the fault that caused last year's devastating earthquake and tsunami off the coast of Japan's Tohoku region. The drilling ship *Chikyu* (pictured) will take temperature measurements near the earthquake's epicentre to measure the heat it generated (see *Nature* 479, 16; 2011).

On 27 April, Japan's Agency for Marine-Earth Science and Technology confirmed that the drill had passed through 6.9 km of water and bored another 860 metres into Earth's crust. This is a record for research drilling in such deep water, surpassing the previous maximum of 7.05 km set by the US vessel *Glomar Challenge* in the Mariana Trench in 1978.

It seems to ignore the effects of the Fukushima disaster, which shuttered the bulk of the country's low-carbon energy supply — its nuclear reactors. Politicians hope to agree on the country's future energy mix by June. See [go.nature.com/z8p7sh](http://go.nature.com/z8p7sh) for more.

### US-China links

US congressman Frank Wolf (Republican, Virginia), who last October tried to ban cooperation between Chinese scientists and the US White House Office of Science and Technology Policy (OSTP), has accepted a compromise that resolves the issue. Under a bill passed by the House appropriations committee on 26 April, OSTP representatives must notify Congress 30 days before meeting with

Chinese officials, and will not meet known human-rights violators. In return, the agency will receive its requested 2013 budget, an increase from previous cuts instigated by Wolf. OSTP director John Holdren is in Beijing this week for annual strategy talks between the two countries.

### Slow down, world

In a policy report on the links between population growth, consumption and human and environmental well-being, the Royal Society in London on 26 April stressed the need for more education and voluntary family planning to stabilize the world's population. The report, *People and the Planet*, says that rich economies need to scale back unsustainable material consumption and

pollution; at the same time, it called for more work to cut extreme poverty.

### Forest-loss fear

Legislation that would relax forest protections in the Brazilian Amazon passed Brazil's lower house of Congress on 25 April, after two years of political wrangling. Environmentalists fear that it will lead to increased deforestation in the Amazon. Eyes are now on the country's president, Dilma Rousseff, to see whether she will use a veto power allowing her to strike sections from the law. See page 19 for more.

### US bioeconomy

US President Barack Obama's administration on 26 April released the *National*

**Bioeconomy Blueprint**, a report outlining how government agencies will boost economic activity based on the biosciences. The document focuses on efforts such as investing in research, translating basic discoveries to industry and reforming regulation. Most of the *Blueprint* simply lists existing initiatives — but the biotechnology industry saw the publication as an encouraging step. See [go.nature.com/gziuke](http://go.nature.com/gziuke) for more.

## Science policy gain

The Union of Concerned Scientists, an advocacy group headquartered in Cambridge, Massachusetts, said on 30 April that it had received a US\$1-million donation from physicist and former IBM vice-president Lewis Branscomb to set up a Centre for Science and Democracy. The centre, to be launched on 17 May, aims to ensure that US policy is based on sound scientific evidence, and to prevent political interference in federal science. See [go.nature.com/njrahd](http://go.nature.com/njrahd) for more.

## Flu work published

The Dutch government is allowing publication of a study that created versions of the H5N1 avian influenza virus capable of spreading between mammals. It granted an export control licence for the work



on 27 April, after the research team's leader, Ron Fouchier at the Erasmus Medical Center in Rotterdam, applied for the licence 'under protest'. Fouchier (**pictured**) had earlier argued that his work should be exempt from export controls, and had said he planned to submit his manuscript without a licence. The study will be published in *Science*; a parallel US-led study is published this week in *Nature* (see pages 5, 7 and 13 for more).

### RESEARCH

## Click for updates

A service that aims to make updates to the research record more clearly visible launched on 27 April. Each year, some 100,000 papers are updated with corrections, retractions or other revisions, but readers may miss these changes. In the 'CrossMark' system, documents have a clickable logo that alerts readers to changes — and that publishers could also use to record other

content, such as a paper's peer-review process and funding sources. CrossMark is optional, but its creator, CrossRef (a non-profit collaborative agency formed by publishers), hopes that it will be widely adopted. See [go.nature.com/jklkdb](http://go.nature.com/jklkdb) for more.

## Drug oversight

A report by the US Institute of Medicine has criticized how the Food and Drug Administration (FDA) manages marketed drugs. The institute told the FDA to produce a user-friendly, publicly available risk-assessment plan for every drug it approves, and to update it throughout the drug's lifetime with details of regulatory actions and safety questions. *Ethical and Scientific Issues in Studying the Safety of Approved Drugs*, published on 1 May, was spurred by revelations in 2007 that the diabetes drug Avandia (rosiglitazone) presented serious heart-attack risks.

### PEOPLE

## Urban science

Physicist Steven Koonin, the former undersecretary for science at the US Department of Energy, will head up an urban science centre in Brooklyn, New York. The Center for Urban Science and Progress, announced on 23 April, is one of three

## COMING UP

### 5 MAY

Japan will shut down its final operating nuclear reactor, at the Tomari power plant on the island of Hokkaido. No reactors taken offline since the Fukushima nuclear accident last year have been restarted.

### 7 MAY

Space-flight firm SpaceX of Hawthorne, California, is scheduled to launch its Dragon capsule in a mission to carry cargo to the International Space Station — a milestone for private space flight. [go.nature.com/nosop7](http://go.nature.com/nosop7)

research centres that will be built on city-owned land, under a regeneration initiative spearheaded by New York City mayor Michael Bloomberg. See [go.nature.com/ymk5ug](http://go.nature.com/ymk5ug) for more.

## EPA adviser

Glenn Paulson, an expert on environmental public health, will be the next chief science adviser to the US Environment Protection Agency, said Lisa Jackson, the agency's head, on 24 April. Paulson has worked at the New Jersey Department of Environmental Protection in Trenton and the Natural Resources Defense Council, based in New York.

### CORRECTION

The story 'US public health' (*Nature* **484**, 146; 2012) incorrectly stated that an Institute of Medicine report urged a doubling of all federal health-care spending. In fact, it urged a doubling of the federal portion of public health spending.

### NATURE.COM

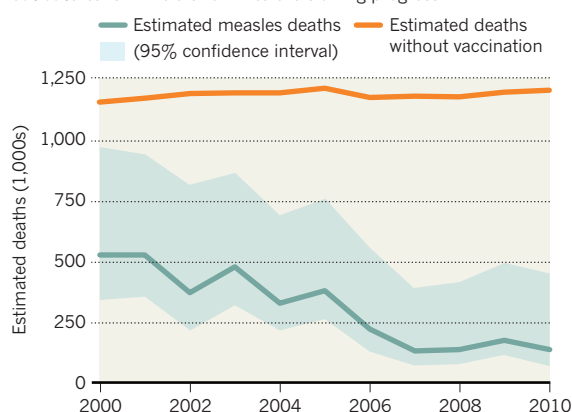
For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Vaccination reduced measles deaths worldwide by almost three-quarters over the past decade, from an estimated 535,300 in 2000 to 139,300 in 2010, says a study published on 24 April (E. Simons *et al. Lancet* <http://doi.org/hvf;2012>). In 2008, the World Health Organization set a target of 90% reduction by 2010. But progress slowed because mass-vaccination campaigns were delayed in India (where 47% of all 2010 measles deaths occurred) and there were measles outbreaks in southern Africa.

## MEASLES DOWN BUT NOT OUT

Deaths from measles decreased 74% in the past decade — but outbreaks in India and Africa are slowing progress.





# NEWS IN FOCUS

**PARTICLE PHYSICS** US neutrino hopes collide with budget realities **p.16**

**CONSERVATION** Brazil moves to dismantle rainforest protections **p.19**

**EGYPT** Nobel laureate's science city caught up in property dispute **p.21**

**HUMAN ORIGINS** Tracing humanity's earliest migrations **p.23**



J. SMALLEY/NATUREPL.COM



A mutant flu virus became more transmissible as it passed between lab ferrets, raising fears that mutations in the wild could create a human pandemic strain.

## VIROLOGY

# Mutant-flu paper published

*Controversial study shows how dangerous forms of avian influenza could evolve in the wild.*

BY ED YONG

Avian H5N1 influenza viruses in the wild may be one small step away from spreading effectively between mammals. That is the sobering message from a controversial study by Yoshihiro Kawaoka at the University of Wisconsin–Madison, published online by *Nature*<sup>1</sup> after months of debate about how to release the findings publicly.

“After wanting to read it for so long, it was like eating again after fasting,” says Vincent Racaniello, a virologist at Columbia University in New York. “And it does not disappoint.”

H5N1, commonly known as bird flu, is highly pathogenic and often lethal in humans, but it cannot spread efficiently between people and cases seem to be rare. To find out if H5N1 could evolve easy transmissibility between humans, Kawaoka and his team mutated its

haemagglutinin (HA) gene, which produces the protein that the virus uses to stick itself to host cells. Because flu viruses in the wild can also gain new properties by swapping genes, the researchers combined this gene with seven others from a highly transmissible flu virus, the H1N1 strain that caused a pandemic in 2009.

Kawaoka found that the hybrid virus could spread between ferrets in separate cages after acquiring just four mutations. Three of these allow the HA protein to stick to receptor molecules on mammalian cells, and the fourth stabilizes the protein. “Before we initiated this experiment, we knew that receptor specificity

is important,” says Kawaoka. “We didn’t know what else was needed.”

Worryingly, some Middle Eastern H5N1 strains can already

recognize human receptors<sup>2</sup>. Kawaoka’s work suggests that they could be just one stabilizing mutation away from being able to spread between humans. Discovering “that HA needs to be stable to be transmissible through the air between mammals” is a key finding, says influenza virologist Wendy Barclay at Imperial College London.

The first hints of Kawaoka’s work emerged last year, along with details of similar experiments led by Ron Fouchier at the Erasmus Medical Center in Rotterdam, the Netherlands. The news sparked intense discussion about whether the benefits of knowing about these potentially dangerous mutations outweighed the risks of publishing them openly. The US National Science Advisory Board for Biosecurity (NSABB) — an independent government advisory board — recommended in December 2011 that both papers should ►

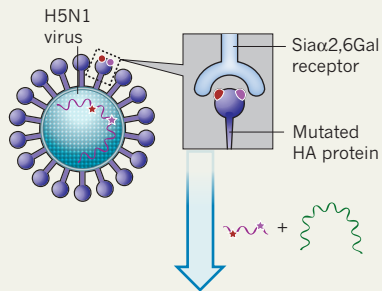
► **NATURE.COM**  
For more, see  
*Nature’s* mutant-flu  
special:  
[go.nature.com/mhmibi](http://go.nature.com/mhmibi)



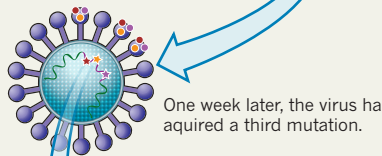
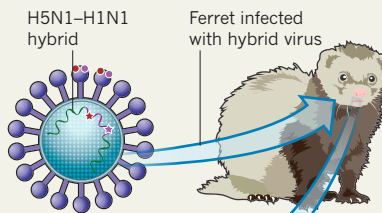
## MAKING MUTANT FLU

Yoshihiro Kawaoka's team randomly mutated the H5N1 virus until a key protein could bind to a mammalian cell receptor. Infecting ferrets with the virus caused additional mutations that made it highly transmissible.

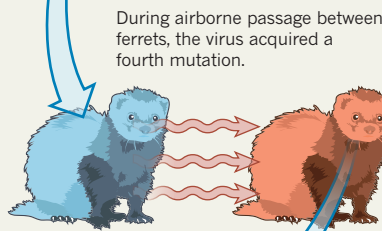
Two mutations in the HA protein allow the virus to stick to receptors on human upper airway cells.



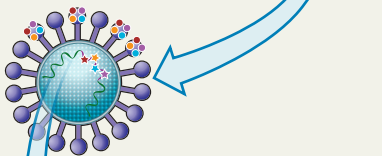
The mutated HA gene was combined with seven genes from the 2009 H1N1 pandemic virus.



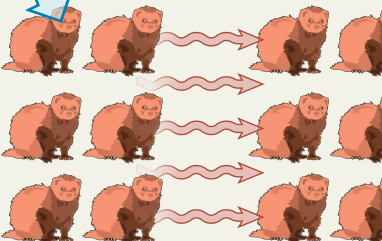
One week later, the virus had acquired a third mutation.



During airborne passage between ferrets, the virus acquired a fourth mutation.



The fourth mutation allowed the virus to spread more effectively, infecting ferrets in neighbouring cages.



► be censored before publication, citing concerns that the strains could be used by bio-terrorists, or that untrammelled proliferation of the work would raise the risk of an accidental release from a lab.

But after a meeting that included international flu experts and health-agency representatives, the NSABB decided in March that revised versions of the two papers should be published in full (see [go.nature.com/5tqfen](http://go.nature.com/5tqfen)). The board was swayed by plans to tighten the oversight of such work, as well as by fresh information about the potential benefits to surveillance. It also acknowledged the difficulties in restricting access to the research. Fouchier has just received an export licence from the Dutch government, which has allowed him to submit his paper to *Science* (see [go.nature.com/2c1hjr](http://go.nature.com/2c1hjr)).

H5N1 viruses prefer to stick to receptor proteins containing Sialic acid (Sia), which are found on the surface of avian cells. In contrast, cells in human upper airways are dotted with receptors containing Sialic acid (Sia), which have a slightly different shape. To see whether H5N1 could come to recognize the human-type receptors, Kawaoka's team randomly mutated its HA protein. Of the 2.1 million different strains they created, only one recognized the human-type receptors instead of the avian ones. This mutant had four amino-acid changes, two of which were necessary for its new-found specificity: one from glutamine to leucine (Q226L) and one from asparagine to lysine (N224K).

Next, Kawaoka fused the mutated HA gene with seven others from the H1N1 2009 pandemic strain. Such a hybrid virus might emerge naturally: both H5N1 and H1N1 have been found in pigs, and Kawaoka had previously shown that they have compatible genes<sup>3</sup>. The hybrid virus evolved further after Kawaoka's team gave it to ferrets — the best animal model for human influenza. After six days, one ferret was riddled with tens of thousands times more virus than the others. On sequencing this particular virus strain, the researchers found that it had acquired a third HA mutation: from asparagine to aspartic acid (N158D).

This new addition allowed the virus to spread between the ferrets for the first time. After the researchers infected animals with the triple-mutant strain, the virus could infect some ferrets in nearby cages by airborne spread. Some of the transmitted viruses had a fourth mutation in HA — a threonine to an isoleucine (T318I) — and these proved even better at spreading in subsequent ferret experiments (see 'Making mutant flu').

The hybrid virus did not kill any of the ferrets it infected. It was slower to spread than the 2009 pandemic H1N1 strain, caused less severe damage in the lungs, and was vulnerable to both the front-line drug Tamiflu and a prototype vaccine against H5N1. It is unclear whether the virus could spread between

humans as well as it did between ferrets, or whether the four HA mutations would confer the same ability on a purely H5N1 virus. The seven H1N1 genes in Kawaoka's hybrid may have contributed to its airborne spread. Nevertheless, "this work reminds us just how vulnerable we potentially are to relatively small changes", says virologist Jeremy Farrar, director of the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam.

The nature of the four mutations suggests that the virus faces an important trade-off as it evolves. Flu viruses need to merge their membranes with those of their host cells in order to release genetic material. The HA protein initiates this process by changing shape in response to the cells' pH, but the three receptor-binding mutations in Kawaoka's new virus trigger this transformation prematurely, reducing the virus's ability to spread. The fourth mutation, T318I, stabilizes the protein, so that it changes shape at the right time. It also makes the virus more resistant to high temperatures. All four mutations are necessary for the virus to recognize the right cell and infect it successfully. "That's the value of our paper," says Kawaoka. "It's not just about identifying mutations. It's about identifying the mechanisms through which they work."

Another group had already found a similar mechanism at work in a hybrid H9N2–H3N2 virus<sup>4</sup>. This virus could spread between ferrets only after acquiring one mutation that allowed it to recognize the right receptors and a second that stabilized the fusion process. The need for specificity and stability may be common requirements for easy transmission.

H5N1 strains in Egypt and some other countries in the Middle East already carry the N158D mutation, as well as other mutations that let the virus bind to human-type receptors. Some also have a mutation in the PB2 gene, which is associated with effective virus replication in mammals. All that is missing is a stabilizing mutation.

Kawaoka says his experiments suggest that rather than watching for specific mutations, it is more important to scan for the traits they bestow. If a virus shows signs of binding to human cells, one further test can show how dangerous it is: heat it to 50 °C to check its stability. "It's a very simple assay," he says.

The study provides "the first clues about what properties of the HA protein, other than receptor specificity, might be important for mammalian airborne transmission", says Racaniello. "It would have been a huge loss not to publish this" ■ SEE EDITORIAL P.5, AND NEWS & VIEWS DOI: [HTTP://DX.DOI.ORG/10.1038/NATURE11192](http://dx.doi.org/10.1038/NATURE11192)

1. Imai, M. *et al.* *Nature* <http://dx.doi.org/10.1038/nature10831> (2012).
2. Watanabe, Y. *et al.* *PLoS Pathog.* **7**, e1002068 (2011).
3. Octaviani, C. P., Ozawa, M., Yamada, S., Goto, H. & Kawaoka, Y. *J. Virol.* **84**, 10918–10922 (2010).
4. Sorrell, E. M., Wan, H., Araya, Y., Song, H. & Perez, D. R. *Proc. Natl Acad. Sci. USA* **106**, 7565–7570 (2009).

## MISCONDUCT

# Trial tests Austrian integrity body

*Unfair-dismissal suit will show whether the young agency's rulings are taken seriously.*

BY ALISON ABBOTT

Austria's fledgling science-integrity agency is facing its first big test, as a protein crystallographer it found guilty of misconduct sues his university for unfair dismissal. If the court upholds the university's move, it would bolster the agency's efforts to change attitudes towards misconduct in the country's universities.

The Austrian Agency for Scientific Integrity was created in 2008 in response to a particularly serious case of misconduct at the Medical University of Innsbruck, where clinical trials of a stem-cell therapy were carried out without ethics approval<sup>1</sup>. By the end of 2011, the agency — an association of universities and research institutes — had dealt with a total of 46 cases, and formally investigated 15 of them. Most involved plagiarism or authorship conflicts, and two involved data manipulation.

The case of crystallographer Robert Schwarzenbacher, which it investigated earlier this year, is by far the most serious yet, says Christoph Kratky, a structural biologist at the University of Graz and chair of the agency's board. It is also the first to have serious consequences for the person accused: Schwarzenbacher lost his job at the University of Salzburg after the agency signalled its conclusion that he was guilty of data fabrication.

The case revolves around a paper co-authored by Schwarzenbacher in the *Journal of Immunology*, which described the molecular structure of a version of a birch-pollen allergen<sup>2</sup>. An article submitted to *Acta Crystallographica* in January pointed out “physically improbable” features in the structure. That article was published in April<sup>3</sup>, along with a response from the original authors saying that “Schwarzenbacher admits to the allegations of data fabrication and deeply apologizes to the co-authors and the scientific community”<sup>4</sup>. However, a note added in proof says that Schwarzenbacher had withdrawn his

admission of fabrication.

When Heinrich Schmidinger, the rector at the University of Salzburg, learned about the concerns at the end of January, he contacted the integrity agency for its opinion on the case. “The case was clear because Schwarzen-



The University of Salzburg will defend in court its dismissal of Robert Schwarzenbacher.

bacher had admitted to fabrication, and the agency's investigation confirmed misconduct,” says Schmidinger, who on 22 March gave Schwarzenbacher three months' notice.

Schmidinger says that he was particularly concerned about preserving the university's reputation. Recalling highly critical reports of the Innsbruck scandal in *Nature*<sup>5</sup>, he says, “I didn't want my university to be referred to in that way in the English-language press.”

Schwarzenbacher now says that the data in question were merely wrong and not fabricated. He is taking the university to court because, he argues, a single case of incorrect data in a publication should not outweigh the benefit to the university's reputation brought by his nearly

100 genuine papers, as well as the large grants that he has won in the past. Schwarzenbacher also claims that “the rector cancelled my contract before the agency reached a conclusion — I cannot see any fairness in [this] conduct”. Schmidinger says that the agency had informed him of its verdict before the dismissal, although it did not issue its formal report until early April.

Austrian professors were once civil servants, with jobs almost guaranteed for life — a situation, still common in many European countries, that can make it very difficult to sack an academic<sup>6</sup>. But since a change in the law in 2002, academics in Austrian universities have had normal contracts that allow the university to dismiss them if they are deemed “unworthy of the confidence of the employer”.

Schwarzenbacher's case, to be heard later this month, will be the first dispute centred on scientific misconduct to come before a labour court. “We'll find out if the courts will accept this type of misconduct as grounds for dismissal, and if they will take the judgements of the agency seriously,” says Kratky.

Although universities are not legally bound to accept the agency's judgements, Kratky says that the agency has “raised consciousness” about scientific misconduct, which has historically been swept under the carpet in Austria. “There is much less cynicism among university rectors” about tackling it, he says, although the agency's formal judgements are not always taken on board by universities.

If the labour court upholds the right of the University of Salzburg to dismiss Schwarzenbacher, other universities will be encouraged to punish misconduct, Kratky says. If it does not, he plans to push for a change to the agency's legal basis, to give its decisions teeth. ■

1. Abbott, A. *Nature* **454**, 922–923 (2008).
2. Zaborsky, N. et al. *J. Immunol.* **184**, 725–735 (2010).
3. Rupp, B. *Acta Crystallogr. F* **68**, 366–376 (2012).
4. Zaborsky, N. et al. *Acta Crystallogr. F* **68**, 377 (2012).
5. *Nature* **454**, 917–918 (2008).
6. Abbott, A. *Nature* **477**, 384 (2011).

  
**MORE  
ONLINE**

## TOP STORY

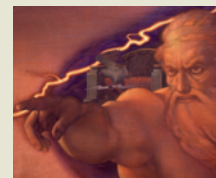


Petition gathers names in bid to get diabetes drug Victoza banned over cancer concerns [go.nature.com/7uu2tu](http://go.nature.com/7uu2tu)

## MORE NEWS

- FDA finds non-compliance with rules on clinical-trial data not to be a major problem [go.nature.com/mruaxb](http://go.nature.com/mruaxb)
- Drug-resistant bacteria go undetected [go.nature.com/mtwzq1](http://go.nature.com/mtwzq1)
- Sunny outlook for space-weather forecasters [go.nature.com/to8j2c](http://go.nature.com/to8j2c)

## MUSE

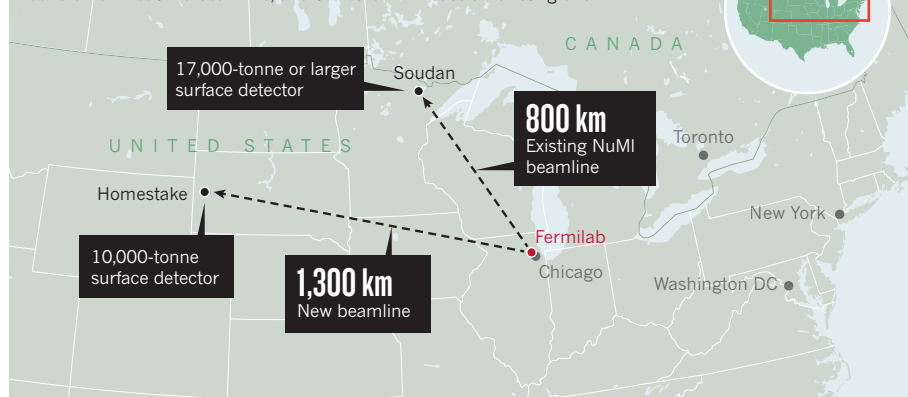


Is it true that rationality is the enemy of religion? [go.nature.com/vatin4](http://go.nature.com/vatin4)



## NEUTRINO FIXES

Two options for saving a planned flagship neutrino experiment replace underground detectors with cheaper surface detectors. One aims to build a new neutrino beamline, and the other will reuse an existing one.



## PARTICLE PHYSICS

# Neutrino project changes focus

*Budgetary constraints force United States to downgrade plans for flagship experiment.*

BY NICOLA JONES

For physicists at the Fermi National Accelerator Laboratory in Batavia, Illinois, the nearly massless particles called neutrinos hold special weight. A proposed project called the Long-Baseline Neutrino Experiment (LBNE) aims to use neutrinos to sort out some puzzles of fundamental physics, and ensure that US particle physics has a centrepiece experiment for the future.

Yet, faced with serious budget constraints, Fermilab is drastically scaling back its plans. It originally intended to fire neutrinos from its base near Chicago at a massive underground detector, filled with 34,000 tonnes of liquid argon, situated 1,300 kilometres away at the Homestake mine in South Dakota. But after a workshop held on 25–26 April, Fermilab is focusing on two cheaper options. Both limit the sensitivity of the experiment and back off from plans for a deep underground detector, ruling out a potential pay-off: watching for proton decay, a prediction of the theory that unifies physics' fundamental forces.

When Fermilab last year shut down the Tevatron, once the world's most powerful accelerator, US particle physicists ceded the 'energy frontier' to

Europe, with its Large Hadron Collider at CERN near Geneva, Switzerland. The LBNE is part of the United States' bid to become a world leader at the 'intensity frontier' — experiments that use beams containing large numbers of particles to study very rare events.

But the original LBNE proposal, estimated to cost up to US\$1.7 billion, was too much for the US Department of Energy to stomach. On 19 March, the department asked Fermilab to reconfigure the experiment into cheaper stages. Fermilab set up a working group, chaired by deputy director Young-Kee Kim, to revise the LBNE proposal. She says that the group aims to keep the cost of the first phase of the project to between \$600 million and \$700 million.

Last week, the working group met with the wider particle-physics community to hash through ways to cut costs while retaining the project's three main science goals: determining the hierarchy of masses of the three known types of neutrino; establishing whether neutrinos disobey 'charge–parity symmetry', which might explain why the Universe contains more matter than antimatter; and searching for proton decay.

The workshop turned out two potentially

affordable options, says Kim (see 'Neutrino fixes'), but both sacrifice some of the scientific goals. One includes a new neutrino beam, but replaces an underground detector at Homestake with a smaller, surface detector that would be filled with 10,000 tonnes of liquid argon. The smaller size means that some statistical power will be lost, and the surface location rules out the possibility of seeing proton decay because particle decays from other sources would confound the results. But Jogesh Pati, a theoretical physicist at the SLAC National Accelerator Laboratory in Menlo Park, California, says that the detector isn't big enough to keep the United States competitive. "With this detector, it would really be withdrawing itself from frontier science," he says.

Proceeding with this option would, however, allow for a larger, or underground, detector to be built in the future, and for upgrades to its beam to support 'Project X' — the next high-intensity project in Fermilab's plans.

The second option is to aim an existing beam, called Neutrinos at the Main Injector, or NuMI, to a 17,000-tonne surface detector 800 kilometres away at the Soudan lab in Minnesota. Because the shorter distance makes it harder to see 'oscillations' between the three types of neutrino, the project would probably need to use data from other experiments to determine the mass hierarchy. Like the first option, this alternative would be unable to watch proton decay. But using an existing beam would be cheaper, and mean that the experiment could get started more quickly.

Some question whether the working group is thinking broadly enough. Neutrino physicist Janet Conrad of the Massachusetts Institute of Technology in Cambridge says that the group could, for example, revive an idea of using water in its detectors, which might be cheaper than liquid argon.

## IFYING FOR THE LEAD

Other projects are already looking at some of the science questions that the LBNE was supposed to crack. Fermilab has one: NovaA (NuMI Off-Axis Electron Neutrino Appearance Experiment), which will start taking data in 2013, aims the NuMI beam at a detector in Ash River, Minnesota. It has a shot at answering the mass-hierarchy question, as does the T2K experiment in Japan. Upgrades planned for Super-Kamiokande in Japan in the next decade should allow it to watch for proton decay. And the HyperKamiokande project, which has yet to be proposed formally, aims to pin down the charge–parity violation issue.

Mel Shochet of the University of Chicago in Illinois, who led the working group's physics team, worries that downgrading the LBNE project will undermine Fermilab's global competitiveness. "It's crucial that we maintain a large facility in the United States," he says. "If we can't find a solution to this, it's a serious problem. But I'm hopeful." ■

**"It's crucial that we maintain a large facility in the United States."**

► **NATURE.COM**  
For more on neutrino experiments, see [go.nature.com/t1gkzb](http://go.nature.com/t1gkzb)



## SECURITY

# US satellite laws under scrutiny

*NASA allegations muddy waters as government recommends reforms to technology–trade rules.*

BY SHARON WEINBERGER

Advocates of international trade and collaboration in space technology thought that they were making headway against rules that restrict both in the name of US security. But on the same day that the US government released a long-awaited report that recommends easing those regulations, allegations surfaced that a NASA director may have broken the rules when he gave foreign nationals access to an agency research facility. It is not yet clear whether the allegations will strengthen the case for preserving current restrictions.

The allegations were brought by unnamed whistle-blowers to US Senator Charles Grassley (Iowa), the ranking Republican member of the Senate Committee on the Judiciary, who asked about the accusations in a letter to NASA administrator Charles Bolden on 18 April. At issue is whether Simon ‘Pete’ Worden, director of the NASA Ames Research Center in Moffett Field, California, gave foreign citizens access to information that falls under the International Traffic in Arms Regulations (ITAR), a set of rules that aims to prevent the transfer of potentially strategic technologies to foreign countries. Critics have long complained that ITAR unduly hampers US companies that seek to export satellite technology; it has also created hurdles for academic and government research institutions — including NASA Ames — that have collaborators in friendly nations such as the United Kingdom and Canada.

Grassley’s office told *Nature* that NASA is investigating the allegations; an agency spokesman confirmed that NASA has referred Grassley’s query to its inspector-general, but declined to comment further. Whatever the outcome, the case is sure to feed a debate reanimated by an 18 April report in which the US Department of Defense and Department of State recommend pulling many non-military communications and remote-sensing satellites out of ITAR’s jurisdiction.

Advocates of export-control reform say that the report is the first encouraging development since 1998, when Congress placed all spacecraft and related equipment under ITAR, following revelations that two US companies had

shared technical information about a launch failure with China, without seeking government approval. The move proved a blow to many US satellite firms, which lost international customers who were unwilling to deal with the licensing rules. Universities with

recalls the legal hoops that the company had to jump through to use a simple stand to hold its Genesis 1 experimental space habitat before the habitat was sent into orbit on a Ukrainian rocket. “This stand was metal, round and had four legs,” says Gold. “Flip it upside down, put a tablecloth on it, and it was indistinguishable from a coffee table.”

To ease such pressures, the report from the defence and state departments urges Congress to shift oversight of many commercial satellites and related activities from the state department, which currently administers ITAR, to the commerce department, where it resided before 1998.

“I think that careful implementation of this report’s recommendations may have the potential to reduce some of the unintended consequences of space export control on universities and university education,” says Kevin Schmadel, vice-president for government relations at the Universities Space Research Association, based in Columbia, Maryland.

But Gold does not think that the recommendations will be put into practice swiftly. “There are hurdles to overcome, and whether the report is enough to build momentum is a question yet to be resolved,” he says. Some lawmakers are likely to oppose any efforts to change the rules, citing national security. Last year, a bill that would have moved regulation of communications satellites back to the commerce department was introduced in the US House of Representatives, but it has not moved forward. And the NASA Ames case could influence opinion, potentially providing a reason not to ease export controls.

Attempts to change ITAR may have to wait until next year, when they could be tacked on to a bill authorizing spending by the defence department. Joel Johnson, an analyst with the Teal Group, an aerospace and defence consulting firm based in Fairfax, Virginia, has long followed the debate over export controls. He says that with federal elections coming up in November, it is unlikely that Republicans in Congress would want to grant a victory on ITAR to President Barack Obama’s administration. “I would be surprised, quite frankly, if anything happens before the election,” says Johnson. ■



**US company Bigelow Aerospace was subject to rigorous export controls when it launched Genesis 1 on a Ukrainian rocket.**

foreign students are also subject to the rules, which cover what can be taught in classes and who can work on satellite-related student projects funded by the government.

Anecdotes abound about the detrimental impact of ITAR. In one case, detailed by the American Association of Universities in Washington DC, a Chinese postdoctoral student at a US university needed a licence to view NASA’s comments on software developed for the agency’s Gravity Probe B project — even though he had written the software himself. And Michael Gold, director of operations in Washington DC for Bigelow Aerospace, a space-technology company based in North Las Vegas, Nevada,



Rural pressure has produced a bill that could see increased deforestation in favour of agriculture in Brazil.

## ENVIRONMENT

# Brazil set to cut forest protection

*Environmentalists pin hopes on presidential veto to reduce harmful impact of weakened legislation.*

BY JEFF TOLLEFSON

The sound of chainsaws in the Amazon rainforest has faded in recent years as deforestation has slowed, last year dropping to less than one-third of its long-term average. But last week, the lower house of Brazil's National Congress passed a bill that observers say could drastically reduce forest protection.

An organized coalition of rural agricultural interests prevailed in vote after vote during debates, approving amendments that would, for example, scale back forest protections along rivers and hills, give state and local governments more authority over forests, and relieve landholders of the responsibility of reforesting illegally cleared land. The bill would also eliminate a requirement that landowners seeking agricultural loans from the government register their land, document any illegal clearance and submit a plan to come into compliance if they have cleared forests illegally.

Environmentalists hope that pressure from conservation groups and media attention on next month's United Nations Conference on Sustainable Development (Rio+20) in Rio de Janeiro will influence Brazil's President Dilma Rousseff to veto the most radical elements of the legislation.

Passed by a 274–184 vote on 25 April, the bill is a revision of the country's 'forest code', which includes a requirement that landowners in Brazil maintain a minimum proportion of native forest on their land, ranging from 20% in the Atlantic forest along the coast to 80% in the Amazon basin.

Although the code has been on the books since 1965, enforcement increased in the past decade under Rousseff's predecessor, Luiz Inácio Lula da Silva. That enforcement, combined with broader agricultural trends and mounting

public pressure, has helped Brazil to reduce deforestation to its lowest level in more than 20 years (see 'Turning the tide').

"Brazil has justifiably held up its record on deforestation in the Amazon in recent years," says Steve Schwartzman, director of tropical forest policy at the Environmental Defense Fund in Washington DC. "But the changes proposed in what passed the House are fully capable of reversing that trend."

The legislation has been working its way through the National Congress since 2010, motivated by a rural backlash against the crackdown on deforestation. The Brazilian Confederation for Agriculture and Livestock, one of the main business groups supporting the bill, argues that the forest code has become too burdensome. Pointing out that around 28% of the country's 851 million hectares is dedicated to agriculture whereas 61% is forest, the organization says that by approving the law, Brazil's Congress has chosen "the path of sustainable agricultural production".

The bill's fate now rests with Rousseff, whom environmentalists have never fully trusted despite her pledges to maintain the policies of her predecessor, including a commitment to an 80% reduction in the average deforestation rate by 2020. Although Rousseff is predicted to sign the bill, many expect her to use her veto powers to remove its most radical provisions, including one that would grant amnesty to people who violated the forest code before July 2008. Rousseff will have 15 days to make her decision once the legislation reaches her desk.

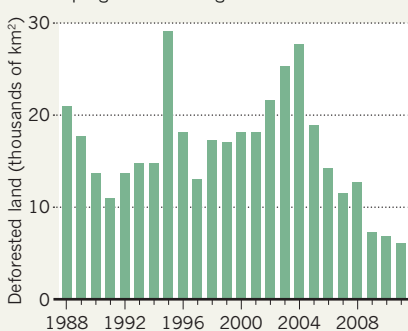
Last December, Brazil's Senate passed a more moderate version of the bill, also opposed by environmentalists, and some senators are now making moves to try to vote again on a fall-back version that Rousseff could accept while rejecting the House bill in its entirety. Some speculate that the administration tacitly allowed the *ruralistas* to advance their agenda in the lower house while planning to veto extreme provisions that were not included in the Senate version, says Paulo Moutinho, an ecologist who heads the Amazon Environmental Research Institute in Brasília. "The bill is so bad now that the president can use her veto and say, 'Look, I'm doing the right thing,'" he adds.

That would help Rousseff burnish her green credentials just in time for Rio+20, the twentieth anniversary of the Earth Summit. She has seen her effigy paraded through the streets of Brasília in recent months, as protesters skewered her record on other environmental issues.

That record includes pushing forward with the controversial Belo Monte dam on the Xingu River in the eastern Amazon, set to be the world's third largest when complete. Opponents of the project accuse the Rousseff administration of disregarding the rights of indigenous peoples and scaling back protected areas. In March, a dozen social and environmental groups warned that Brazil is in danger of being "both host and villain" at Rio+20. ■

## TURNING THE TIDE

Critics fear that changes to Brazil's forest code will end progress in slowing deforestation.



SOURCE: INPE



# Pre-term births on the rise

*But simple measures could cut the mortality rate of premature babies in poorer countries.*

BY EUGENIE SAMUEL REICH

Premature birth is the biggest cause of infant mortality worldwide, yet most of those deaths could be readily prevented, according to a 2 May report from the World Health Organization (WHO) and child-health advocacy groups. The report summarizes the results of a comprehensive survey of the problem, and says that pre-term births are on the rise—a worrying trend—but that low-income countries could reduce deaths among these infants by introducing a few affordable key health-care practices.

Premature babies, defined as those born before 37 weeks' gestation, have higher rates of death and serious illness in the days and weeks after birth than do babies born at term, and are also more likely to be left with life-long disabilities such as blindness and cerebral palsy. The report discusses estimates of the pre-term birth rates and associated infant mortality for most of the countries in the world; the data, which cover 184 countries, are presented in two papers under review at *The Lancet*.

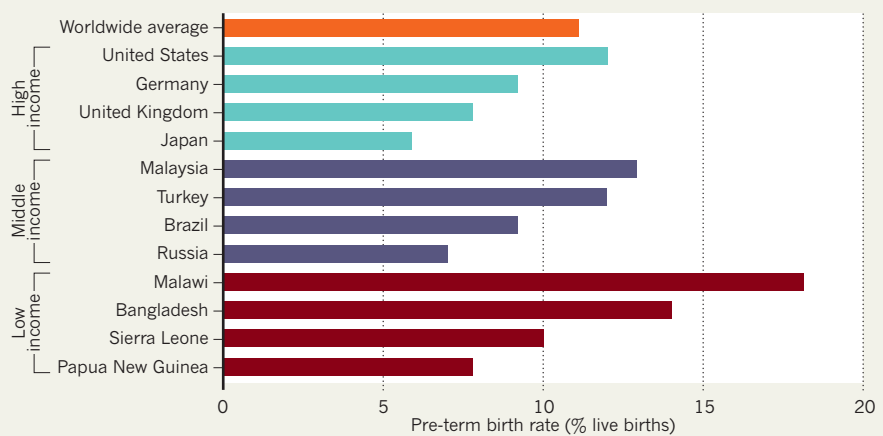
This is the first detailed breakdown and comparison of premature birth rates country by country, compiled from records at national registries and health-care centres, and applying consistent definitions, says Elizabeth Mason, a public-health specialist at the WHO in Geneva, Switzerland, who is a co-author of the report. The researchers hope that their three-year study will encourage countries to take steps to improve survival rates. "We use the evidence to derive policy and action. We don't believe in research for research's sake," Mason says.

The survey estimated that 15 million premature babies are born each year, of which more than 1 million will die. Pre-term births range from a low of 5–6% of live births in Japan and some European countries, including Latvia and Sweden, to more than 16% in many African countries, with a high of 18.1% in Malawi (see 'Ahead of schedule'). Some high-income countries have surprisingly high rates, such as the United States with 12%. Time-trend data compiled for 65 countries, mostly with high incomes, show that rates are rising in most places, the report says, although the underlying causes differ.

***"An African woman cares just as much if she's lost her baby, but her voice hasn't been heard."***

## AHEAD OF SCHEDULE

Figures for a selection of countries surveyed by the World Health Organization show that pre-term birth rates vary widely within and across income groups. Rates are increasing in most countries.



In high-income countries, the causes of premature birth are thought to include mothers delaying childbirth to later ages, the growing use of fertility treatments, which can produce multiple pregnancies, and an increase in obesity and diabetes. Pre-term births by caesarean section, both those that are medically warranted and those done for convenience, also contribute to the increase. In many low-income countries, infections, including malaria and HIV, seem to be at least partly responsible for the increase in premature labour. But much of the trend, especially in Africa, remains unexplained, and more research is needed into the causes, says Mason.

The new data reveal striking disparities in mortality rates across different countries, says Chris Howson, an epidemiologist who co-authored the report on behalf of the child-health advocacy group the March of Dimes Foundation in White Plains, New York. In high-income countries almost 95% of babies born between 28 and 32 weeks will survive, but in low-income countries 70% will die.

The mistaken assumption by health-care workers in low-income countries that nothing can be done for premature babies is one probable reason for the survival gap, says Melissa Gladstone, a paediatrician at the University of Liverpool, UK, who has studied the problem extensively in Malawi. She agrees that affordable countermeasures could cut the death rate due to premature birth.

One such practice

is 'kangaroo' mother care, in which premature babies are strapped skin-to-skin to their mothers for 24 hours a day, keeping them warm without an incubator. In addition, steroid injections for mothers entering premature labour, which trigger a fetus's lungs to mature rapidly, cost about US\$1 and can be given at a rural clinic. The report argues that most of the reduction in neonatal mortality in the United States and United Kingdom over the past 70 years resulted from better obstetric and neonatal care and pre-dated the introduction of expensive neonatal intensive-care units.

Yet many poor countries will struggle to introduce even such simple measures, says Gladstone. "Staffing and motivation is a huge problem," she says. Underpaid nurses, often caring for dozens of children, may lack the time and inclination to explain and supervise kangaroo mother care, for example. Enis Baris, a public-health specialist at the World Bank who has studied improvements in neonatal outcomes achieved by Turkey, a middle-income country, says that implementing such measures will take strong leadership by government officials.

The lesson from developed countries is that governments tend to act in response to the advocacy of bereaved parents, says Joy Lawn, a paediatrician at Save the Children in Cape Town, South Africa, who points to the success in raising awareness of parent-founded groups such as the March of Dimes in the United States and Tommy's in the United Kingdom. "An African woman cares just as much if she's lost her baby, but her voice hasn't been heard," Lawn says. ■

➔ **NATURE.COM**  
For more on ways of improving child health see:  
[go.nature.com/uuijx](http://go.nature.com/uuijx)

SOURCE: H. BLENKOWE ET AL FOR THE WHO





Nile University and Zewail City of Science and Technology both claim ownership of the same buildings.

## HIGHER EDUCATION

# Universities clash by the Nile

*Property dispute dogs Egypt's plans for a science city.*

BY KATHARINE SANDERSON

**A**n ambitious scheme to create Egypt's first science city is finally becoming a reality. Pristine glass and steel buildings are waiting for researchers and students just outside Cairo, and a grand opening is planned for September.

But a bitter property dispute is overshadowing enthusiasm for the technology park and university complex. In a quarrel that has roots in the political turmoil in Egypt in recent years, another university says that it owns some of the buildings, and that its research efforts are being thwarted.

The current occupant of the site is the Zewail City of Science and Technology, spearheaded by Ahmed Zewail, a Nobel prizewinning Egyptian chemist at the California Institute of Technology in Pasadena. Zewail first proposed the concept in 1999; a cornerstone was laid the following year, but the project was soon stalled by political instability and bureaucracy, says Zewail.

The government then allocated part of the site to Nile University,

established in 2006 as Egypt's first research university by the non-profit Egyptian Foundation for Technology Education (EFTED). The institution now has about 370 students, three-quarters of whom are postgraduates.

Nile University began construction work at the site, and research staff had just moved into their new labs and offices when the January 2011 revolution struck. In the aftermath of the uprising, many of the ousted regime's decisions were overruled and reversed, including its gift of land to Nile University, which tainted the university — unfairly, staff say — through its connection to the old government. In February last year, the board of EFTED signed Nile University's buildings and land over to the interim government, and staff moved out. "That was under pressure from the government, and it was illegal," says Moustafa Ghanem, vice-president for research at Nile University and a research fellow in computing at Imperial College London.

In May 2011, the government reassigned the land and buildings to Zewail City of Science and Technology and declared the complex a National Project of Scientific Renaissance. Since then, with nowhere suitable to work, Nile University researchers have been forced

to rent labs that are available for only limited hours, says Mohamed Abdel-Mottaleb, director of the university's nanotechnology programme. As a result, research is "going at a snail's pace", he adds.

Zewail had several meetings with representatives of Nile University last year to try to resolve the situation, but talks — and goodwill — have broken down. "We actually signed an agreement" to merge Nile University with the science city, says Zewail. But Nile University agreed to the deal only on the condition that it keep an independent identity, says Ghanem. Because that has not yet been guaranteed, he says, the deal is off.

The government has offered to compensate Nile University for the money it spent kitting out the buildings and labs, but without the buildings, "what would we do with that money?" asks Ghanem. Unless the situation is resolved, the university's future remains precarious, he adds. "Nile University is now surviving hand to mouth. We'll survive one more term."

Nile University students have launched a vigorous Internet campaign to save their university, using Twitter, Facebook and YouTube to gather support. Meanwhile, Zewail is pushing ahead with plans for his university, promising that curricula for undergraduate courses will be online "within weeks". "It is not just a university," he says, "we are trying to bring some pride back to Egypt, showing that it is capable of doing research and science at the world level."

A government committee heard Nile University's case earlier this month, and on 29 April decided that it would bring together the two parties to find a compromise. "We are trying to work with Dr Zewail and Nile University to reach a cordial agreement," says Abdel Tawab Othman, a member of the committee and a philosopher at Al-Azhar University, Cairo. Several options are being considered, says Othman: "Can Nile get its buildings back and Zewail get a different piece of land? Or would they agree to compensation?"

Zewail's position is that the government has legally given the buildings to his science city, and it is now too late to change that. He says that Nile University should accept the government's offer of compensation.

Ghanem hopes that Egypt's presidential elections at the end of this month will help to break the impasse. "We hope that once one of them is sworn in, we should see some action taken," he says. ■

*Additional reporting by Mohammed Yahia*

### CORRECTION

The News Feature 'Good science bad science' (*Nature* **484**, 432–434; 2012) wrongly located David Keith at the University of Calgary. In fact, he is at Harvard University.

➔ **NATURE.COM**  
For more on the Arab Spring and its impact on science, visit:  
[go.nature.com/ox5oag](http://go.nature.com/ox5oag)

# PEOPLING THE PLANET

**N**ot long ago, the story was simple. A vanguard of modern humans left their African birthplace 50,000–60,000 years ago and quickly conquered Asia. They turned left into Europe some 40,000 years ago, later crossing the Bering Strait and marching southward into the Americas. With their advance, Neanderthals and other earlier peoples dwindled and vanished.

But in the past five years, the picture has grown more complex — and more interesting.

Few question the idea that modern humans are all emigrants from Africa. But when their journey began, when it ended and what they did along the way makes for a deepening mystery, explored in this issue of *Nature*.

Discoveries on the Arabian peninsula, for example, show that modern humans were camped on the doorstep of Asia more than 100,000 years ago, nearly twice as long ago as anyone thought. If they went farther at that early date — and some archaeologists insist that they must have — their presence would explain a smattering of ambiguous artefacts and fossils around Asia (see page 24).

Elsewhere, humans definitely arrived ahead of schedule. Sensitive new radiocarbon-dating techniques show that the first modern humans reached Europe thousands of years earlier than was thought, implying a lengthy coexistence with Neanderthals there (see page 27). And the picture of big-game hunters following an inland route from Asia to the Americas 13,000 years ago has been obliterated by a barrage of reports of older sites. Archaeologists are studying DNA, ancient and modern, for clues to when and how the first Americans arrived (see page 30).

The most dramatic change, however, concerns the archaic peoples whose world we inherited. In the past two years, ancient-DNA researchers have deciphered the full genome sequences of Neanderthals and a hitherto unknown group called Denisovans, then compared them with modern human genomes. The startling upshot: genetic traces of our vanished cousins live on in people today (see page 33). Just where and how the ancient trysts took place is yet to be revealed, as researchers continue to unravel the human story. ■

SEE EDITORIAL P.6



## PEOPLING THE PLANET

Interactive map of migrations:  
[go.nature.com/nmz8gd](http://go.nature.com/nmz8gd)







# EASTERN ODYSSEY

BY TIM APPENZELLER

*Humans had spread across Asia by 50,000 years ago. Everything else about our original exodus from Africa is up for debate.*



One day some 74,000 years ago, in a swampy valley in the south of India, dawn never came. In the half-light, greyish dust sifted down, blanketing the ground and turning trees to ghosts. Far to the east, a volcano called Toba on the Indonesian island of Sumatra had unleashed one of the greatest eruptions ever known, flinging thousands of cubic kilometres of rock into the atmosphere and spreading a pall of ash across southern Asia.

Clive Oppenheimer, a volcanologist at the University of Cambridge, UK, has studied the ash deposits in India's Jurreru Valley to reconstruct the events that followed. Within days, the trees shed their whitened foliage; rains later swept ash into layers several metres thick on the valley floor. Eventually, the lakes and swamps vanished, perhaps because

The Andaman Islands: a stop for migrants 'beach-hopping' around Asia?

the climate had become drier and cooler. Toba had transformed a lush habitat into a wasteland<sup>1</sup>.

The catastrophe had witnesses. Archaeologists digging beneath the ash layer have found stone artefacts indicating that humans were living in the valley before the eruption. But were they modern humans — people like us — or some other, now extinct, branch of the human lineage?

Today, the thick ash deposits of the Jurreru Valley mark a division not only in the geological record but also between archaeologists debating one of the field's biggest questions: when and how did modern humans leave their African cradle and colonize Asia, Earth's largest landmass? It was the first great expansion of the human species, carrying people some 12,000 kilometres to Australia by about 50,000 years ago. But just how early the pioneers set out is controversial — as are the routes they followed, the tools they carried and, most fundamental of all, what triggered the migration. Were they enticed into the wider world by a favourable climate, or propelled by a revolution in technology and culture?

In archaeologists' shorthand, the debate boils down to a simple question: pre-Toba or post-Toba? At one pole of the debate, Paul Mellars at the University of Cambridge argues passionately that modern humans left Africa long after the Toba eruption, 60,000 years ago at the earliest. Equipped with new technologies, including bows and arrows, they beach-hopped along the coastline of the Arabian peninsula, India and southeast Asia, reaching Australia in short order. Genetic analyses of contemporary Asians that point to a late, rapid colonization have bolstered his confidence. "I'm more convinced than ever that I'm right," he says, before adding, "I guess they all say that."

His opposite number, Michael Petraglia at the University of Oxford, UK, certainly does. He is convinced



## PEOPLING THE PLANET

Interactive map of migrations:  
[go.nature.com/nmz8gd](http://go.nature.com/nmz8gd)

O. BLAISE/GETTY



that people spread into Asia at least 74,000 years ago, and perhaps as early as 125,000 years ago, well before Toba, during a wet, warm interlude between ice ages, carrying tools no more sophisticated than those made by earlier humans. Rather than following the coast, he thinks they wandered along river valleys and lake shores, advancing and retreating as the environment allowed, more like wildlife than a wave of colonists (see 'Asian migration'). He takes heart from discoveries that testify to the presence of modern humans in the Arabian Peninsula, on the very doorstep of Asia, more than 100,000 years ago. "I'm feeling more and more confident through time," he says.

It is a full-throated academic duel. Petraglia calls Mellars' view "baseless archaeologically"; Mellars vows to "demolish the pre-Toba model". Their passion is fuelled both by the prize of understanding the first great human migration — and by the lack of decisive evidence. That leaves many other archaeologists in the same position as Chris Stringer of London's Natural History Museum, who says he is "sitting on the fence" until researchers unearth more data. The wait may not be long. Archaeologists are busily hunting for artefacts and even fossils of the first modern Asians, which could finally put the debate to rest.

### COASTAL EXPRESS

Most of Asia is "a blank map, if you will, in its prehistory", says Petraglia. Scattered fossils record the archaic peoples — *Homo erectus*, Neanderthals and the recently identified Denisovans — who had the continent to themselves before modern humans moved in. But few artefacts and no convincing fossils record this arrival. Researchers have mostly relied on the DNA of people today to reconstruct the ancient story.

Geneticists collected mitochondrial DNA (mtDNA) — which is inherited only from the mother — across much of Asia, focusing on isolated native groups thought to be descendants of early human settlers in their areas. They identified distinctive mtDNA variants, or haplotypes, and compared them to create a family tree of the African exodus. To date its roots and branches, they used estimates of mutation rates as a molecular clock.

The tree's root is marked by a haplotype called L3 that originated before humans left Africa; its genetic signature is found in many Africans and every non-African today. The latest readings of the molecular clock date L3 to 60,000–70,000 years ago<sup>2</sup>, suggesting that humanity left

Africa a few thousand years after Toba. The three next-oldest haplotypes — immediate descendants of L3 outside Africa — are 60,000–65,000 years old. All three gave rise to multiple 'starbursts' of genetic variation, scattered all the way from Arabia to Bali. "For that to happen," says geneticist Stephen Oppenheimer of the University of Oxford, "people would have had to move very fast, before new mutations occurred."

The most plausible route for a rapid migration is along the coast of the Indian Ocean, which could have been followed for thousands of kilometres without the need to master new environments. By the

**"PEOPLE SAY MUTE STONES SPEAK. THEY DON'T. WE'RE THE ONES WHO IMPOSE OUR VIEWS ON THEM."**

mid-2000s, most researchers accepted the 'coastal express', as this post-Toba migration scenario is sometimes known. Later analyses of mtDNA and of the male Y chromosome suggested that the exit from Africa was even more recent, perhaps less than 60,000 years ago. And in an influential 2006 review in *Science*<sup>3</sup>, Mellars argued that the sparse archaeological record of Asia and Australia not only supported the genetic evidence of a rapid post-Toba migration, but showed how cultural advances could have helped to drive it.

There is no trace of any ancient beachcombers or their camps, which Mellars says is no surprise because sea-level rise since the most recent ice age has flooded the former coastline. "Your earliest sites in any area are going to be under the sea, and 20 kilometres or more out to sea," he says.

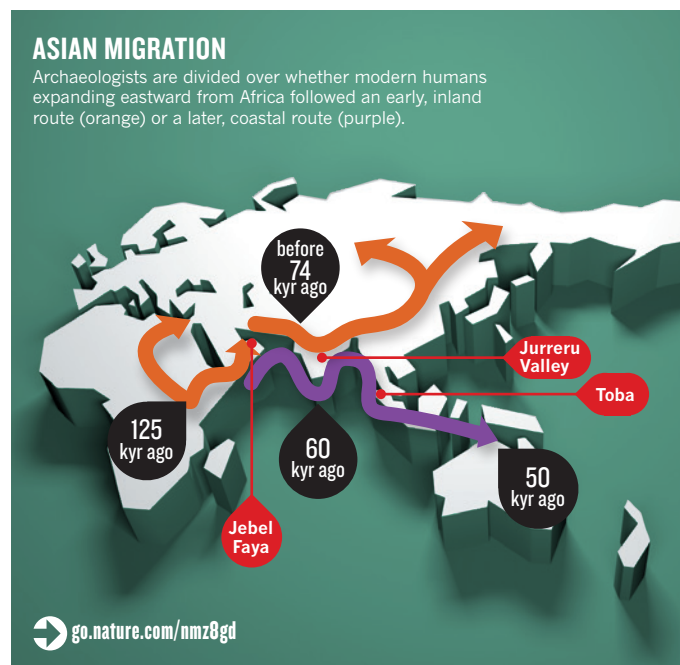
But Mellars sees important clues further inland. The earliest artefacts from the Jurreru Valley and other sites in India and Sri Lanka are simple stone points, scrapers and cores — the lumps of rock left after points were flaked off. All could have been made by archaic humans. Later, well after Toba, they give way to much more sophisticated handiwork: small, finely worked 'microlithic' blades that might have served as arrowheads, shaped bone points, beads and pieces of ostrich eggshell adorned, in one case, with cross-hatching.

The change in the record "appears almost overnight", Mellars says, and mirrors the transformation seen in Europe some 45,000 years ago, when modern humans arrived, bringing sophisticated tools and ornaments that outshone the cruder handiwork of the continent's resident Neanderthals. If Europe's great cultural flowering commenced with a wave of modern human immigrants, he asks, why should Asia's be any different?

Mellars suggests that the migrants carried technology from a culture known as Howiesons Poort, which flourished in southern and eastern Africa from 65,000 to 50,000 years ago. This sophisticated technology, he thinks, was key to launching modern humans out of their native continent. "Innovation would have enormously facilitated expansion eastward," he says.

Critics say the theory faces a serious obstacle, however. Although the genetics suggest that the coastal express dispersed people across Asia at least 55,000 years ago, the earliest microlithic tools found so far in South Asia are no more than 40,000 years old. Mellars is unfazed by the lag, noting that sea-level rise would have flooded the coastal pioneers' earliest sites. The Asian locations with microblades should be younger, he says — because they lie well inland. "That's a totally different environment. You've got a lot of adapting to do, and you won't do it overnight."

While Mellars was picturing a late and rapid expansion of culturally advanced people, Petraglia was digging beneath the grey ash beds in the Jurreru Valley and speculating about a very different scenario. He and his colleagues had excavated simple scrapers and cores from below the ash, dated to between about 77,000 and 74,000 years ago<sup>4</sup>. Neanderthals or other archaic humans could have made them — but, says Petraglia, "no one has ever argued for Neanderthals in India, ever". Instead,



he and his colleagues argued, the artefacts were reminiscent of ones from southern Africa, where the only toolmakers were unquestionably modern humans.

To Petraglia, the implication was clear: modern humans found their way to India before Toba blew its top. As for the microblade technologies that Mellars attributes to a wave of later migrants, Petraglia argues that a long-established population of modern humans developed the new tools on their own.

He dismisses some geneticists' conclusion that clues in modern human DNA rule out a migration from Africa to Asia any earlier than 60,000 years ago. "My issue with these guys is, what are they sampling? They're sampling modern humans that live today — a small subset of what out-of-Africa was." The genetic signal of earlier arrivals has simply been lost, he says, as populations shifted and vanished.

Petraglia's claim intrigued at least one geneticist. Stephen Oppenheimer studies mtDNA and contributed to the coastal-express model<sup>5</sup>. In his view, the migration had to have been rapid and coastal — but not necessarily as recent as other geneticists have insisted. A clue to an earlier date, he says, comes from mtDNA studies in India. One set of variants was less diverse than expected, suggesting to him that the first modern humans in India suffered some kind of catastrophe that reduced their numbers to almost nothing.

The cause, he speculated, might have been the Toba ash cloud, which would imply that modern humans were already established in India at the time of the eruption. He was also intrigued by other hints of a pre-Toba population of modern humans in Asia — such as a foot bone from the Philippines and a skull from China — although all were of uncertain age or origin.

Oppenheimer notes that DNA is an inexact clock. He concedes that the most likely dates for the L3 haplotype — and the exit from Africa — fall after Toba. But "the logic is in the error bars", he says. "The upper limit for the most recent confidence estimate is 79,000 years ago, which means the L3 date straddles Toba." If DNA can't resolve the dispute, he says, the clinching evidence will have to come from archaeology.

## GREEN DESERT

Petraglia and his supporters say that evidence has grown much stronger in the past two years, thanks to the discoveries that put modern humans in Arabia more than 100,000 years ago. What he calls the "smoking gun" comes from Oman, where a group led by Jeffrey Rose of the University of Birmingham, UK, found 106,000-year-old stone relics that are as diagnostic of modern humans as a lost mobile phone would be today. Consisting of distinctive triangular cores and long spear points made from them, this kind of technology was first excavated in Nubia, a region in northern Sudan, at sites inhabited by modern humans. Yet here they were, more than 2,000 kilometres away, on the other side of the Red Sea<sup>6</sup>. "There's no question," says Rose. "It's the same people."

They reached Oman not with the aid of superior technology, Rose says, but because climate and ecosystems favoured the move. At the time, the Arabian peninsula was a moist savannah teeming with game. "The territory next door suddenly turned green and people from north-eastern Africa moved in," says Rose. "These people would have been moving into what they already knew."

But did these early explorers press on, from Arabia into Iran and India? No, says Mellars. Theirs was a failed expansion, like one

documented in Israel, where modern humans were present more than 100,000 years ago but then vanished. When the climate cooled and dried about 75,000 years ago, turning Arabia back into a desert, the Nubian pioneers either died out or retreated to Africa. "There's not a smell, not a whiff [of Nubian technologies] that has ever been detected in India," says Mellars. "If Mike Petraglia could come to me with one of those Nubian cores and say, 'Look, we found this in India,' I would get down on my knees and say, 'Sorry, Mike, I got it wrong.'"

Even if the Nubian toolmakers did not wander farther into Asia, other

ancient Arabian populations might have, Petraglia and others say. At a site in the United Arab Emirates called Jebel Faya, Simon Armitage of Royal Holloway, University of London, and his colleagues have found even older artefacts, dating to as early as 125,000 years ago and resembling objects made by modern humans in eastern Africa<sup>7</sup>.

Jebel Faya sits on a promontory jutting toward Iran, and its inhabitants could easily have pushed eastward a few thousand years later. By 110,000 years ago, ice sheets were beginning to build up far to the north and sea level was dropping, markedly narrowing the Straits of Hormuz and easing the crossing to Iran. "It's extremely plausible that a population at Faya could have moved on," Armitage says.

But plausibility is perhaps the strongest argument that Armitage and others favouring the pre-Toba model can make. If a band of pioneers did wander from Arabia into a welcoming green Asia during the last interglacial,

their tracks are faint indeed.

The pre-Toba artefacts from the Jurreru Valley look nothing like the Arabian ones, says Anthony Marks of Southern Methodist University in Dallas, Texas, who studied the Jebel Faya material. And the archaeologist who analysed the oldest relics from the Jurreru Valley and provided key support for the claim that they are the handiwork of modern humans is no longer so sure. Chris Clarkson of the University of Queensland in Brisbane, Australia, a frequent collaborator of Petraglia's, now thinks they might be the work of an unidentified population of archaic people.

Clarkson and others say it is simply too soon to know for sure whether our ancestors were in India to watch the volcanic ash rain down 74,000 years ago. Pre- or post-Toba: either scenario rests on sparse and ambiguous artefacts. "People say mute stones speak," says Marks. "They don't. They just lie there. We're the ones who impose our views on them."

What is needed is what archaeologists always need: more and better evidence. "We're at the tip of the iceberg, really," says Petraglia. "We've done the best we can with a few sites." On the Arabian Peninsula, along the coast of the Indian Ocean, and in the heart of India, archaeologists continue to search. Perhaps they will find tools clearly made by modern humans before Toba, or a coastal encampment left by later colonists, littered with microlithic blades. Or, most definitive of all, a skull entombed beneath the Toba ash, silently announcing the presence of modern humans — or ceding the ground to our archaic cousins. ■

**Tim Appenzeller** is chief magazine editor for Nature.

1. Oppenheimer, C. *Eruptions that Shook the World* (Cambridge Univ. Press, 2011).
2. Soares, P. et al. *Mol. Biol. Evol.* **29**, 915–927 (2012).
3. Mellars, P. *Science* **313**, 796–800 (2006).
4. Petraglia, M. et al. *Science* **317**, 114–116 (2007).
5. Macaulay, V. et al. *Science* **308**, 1034–1036 (2005).
6. Rose, J. I. et al. *PLoS ONE* **6**, e28239 (2011).
7. Armitage, S. J. et al. *Science* **331**, 453–456 (2011).



Tools such as this unfinished hand axe found at Jebel Faya on the Arabian peninsula suggest that modern humans arrived there more than 125,000 years ago.

H.-P. UERPMANN



# DATE WITH HISTORY

BY EWEN CALLAWAY



*By revamping radiocarbon dating, Tom Higham is painting a new picture of humans' arrival in Europe.*

**B**eside a slab of trilobites, in a quiet corner of Britain's Oxford University Museum of Natural History, lies a collection of ochre-tinted human bones known as the Red Lady of Paviland. In 1823, palaeontologist William Buckland painstakingly removed the fossils from a cave in Wales, and discovered ivory rods, shell beads and other ornaments in the vicinity. He concluded that they belonged to a Roman-era witch or prostitute.

"He did a good job of excavating, but he interpreted it totally wrong," says Tom Higham, a 46-year-old archaeological scientist at the University of Oxford's Radiocarbon Accelerator Unit. Buckland's immediate successors did a little better. They determined that the Red Lady was in fact a man, and that the ornaments resembled those found at much older sites in continental Europe. Then, in the twentieth century, carbon dating found the bones to be about 22,000 years old<sup>1</sup> and, later, 30,000 years old<sup>2</sup> — even though much of Britain was encased in ice and seemingly uninhabitable for part of that time. When Higham eventually got the bones, his team came up with a more likely scenario: they were closer to 33,000 years old and one of the earliest examples of ceremonial burial in Western Europe.

"It is another sobering example of cocked-up dates," says Higham, whose laboratory is leading a revolution in radiocarbon dating. By developing techniques that strip ancient samples of impurities, he and his team have established more accurate ages for the remains from dozens of archaeological sites. In the process, Higham is rewriting European history for around 30,000–50,000 years ago — a time referred

to as the Middle-to-Upper Palaeolithic transition — when the first modern-looking humans arrived from Africa and the last Neanderthals vanished. Higham thinks that better carbon dating will help to resolve debates about whether the two ever met, swapped ideas or even had sex. It might even explain why humans survived and Neanderthals did not.

"I admire him," says Paul Mellars, an archaeologist from the University of Cambridge, UK, and an expert on this period in Europe, for "the sheer doggedness and sense of vision" he has for improving radiocarbon dating of the Palaeolithic. That vision sometimes clashes with other scientists' views, but Higham makes no apologies for his interpretations as long as the dates are solid. "I want to know the truth" is something he says a lot.

## A WOOLLY FIELD

If you Google 'archaeologist' and 'Higham', the first hit is likely to be Charles Higham, a 72-year-old professor who has charted the origins of agriculture and government in southeast Asia. Tom was born in Cambridge, where his father was based until 1966. Charles then moved the family and nine-month-old Tom to New Zealand's rugged south island to start an archaeology department at the University of Otago in Dunedin. As a teenager, Tom spent summers at Ban Na Di, a study site in northeast-

ern Thailand, where his duties included helping with human excavations and brewing tea for the crew.



## PEOPLING THE PLANET

Interactive map of migrations:  
[go.nature.com/nmz8gd](http://go.nature.com/nmz8gd)



Tom didn't originally plan to follow his father's path. As a child he was obsessed with the history of the American West. At university, he planned to study geography and glaciology, but switched to archaeology after excelling in an introductory course taught by his father that he had signed up for on a whim. But his enthusiasm soon waned. "I got less and less interested in archaeology because it was so subjective and woolly."

The reasons for that woolliness were partly technical and partly historical, dating back to before the Highams' time. Archaeology before carbon dating relied on two principles: older things are buried beneath younger things, and people with cultural ties make similar-looking objects, such as stone tools. But dates were hard to come by. In the early nineteenth century, the Danish historian Rasmus Nyerup wrote that most of early human history was "wrapped in a thick fog"<sup>3</sup>. "We know that it is older than Christendom," he wrote, "but whether by a couple of years or a couple of centuries or even by more than a millennium, we can do no more than guess."

The fog began to lift in the middle of the twentieth century, when US chemist Willard Libby and his colleagues<sup>4</sup> showed that all formerly living things bear a clock powered by radioactive carbon-14. Organisms incorporate tiny amounts of this isotope as they grow, and they maintain a constant ratio between it and other, non-radioactive, carbon isotopes throughout their lives. After death, the carbon-14 decays with a half-life of about 5,730 years, and the dwindling ratio serves as a time stamp. Libby's team proved the accuracy of this 'clock' on objects of known age, such as Egyptian mummy tombs, and bread from a house in Pompeii, Italy, that was burned during the eruption of Vesuvius. Libby earned the 1960 Nobel Prize in Chemistry for his work.

The clock gets less accurate as the samples age, however; cruelly, it begins to fail at one of the most interesting times of human history in Europe. Within 30,000 years, 98% of the already vanishingly small quantities of carbon-14 in bone is gone. And carbon-14 molecules from surrounding soil start to seep into the fossils. Collagen, the part of bone that contains the most carbon suitable for dating, sops up contaminants like a sponge, creating a false record. If just 2% of the carbon atoms are contemporary, then a 44,000-year-old bone will return a carbon date of 33,000 years old, Higham calculates.

Most of the thousands of carbon dates from archaeological sites from the Middle-to-Upper Palaeolithic era are wrong, say scientists, perhaps even as many as 90%. As a result, archaeologists can agree on the history of this era only in the broadest of brushstrokes.

Tom found himself drawn to the quantitative side of archaeology to help fill in those details. His father had counselled that if he wanted a future in the field, Tom ought to join the push to make it a more rigorous science, emphasizing testable theory, experiment and statistics. So, at his father's urging, Tom applied for and completed a PhD at the University of Waikato's Radiocarbon Dating Laboratory in Wellington, then did a postdoc there. And when a faculty position became available at a better-funded lab at the University of Oxford in 2000, he moved back to his birth country.

Any idea that archaeology hasn't gone in the direction that Charles predicted is dispelled by a visit to his son's workplace. Its centrepiece is a giant £2.5-million (US\$4-million) particle accelerator, which is used to tot up the number of radioactive carbon molecules in a sample.

Similar machines have been used for carbon dating since the 1970s and have allowed scientists to date smaller samples with more precision than before. But they have also produced their share of erroneous dates. "People used to take bones, grind them up and date them, and you got all kinds of dates because no one bothered to check if there was collagen or not," says Ofer Bar-Yosef, an archaeologist at Harvard University in Cambridge, Massachusetts. And rather than damage valuable human bones or animal bones marked with cuts from stone tools, scientists tended to date fragments of unidentified animal bones found alongside human remains, assuming, not always correctly, that they coincided with human occupation. "It just breaks your heart to see what people

have dated before. They've basically dated pieces of shit," Higham says.

His team didn't change the machine — the secret to more accurate dating lies in the rigorous way the samples are processed beforehand. The team typically starts with bones that are linked unequivocally with human occupation, such as cut-marked bones. To remove contaminants such as decayed organic matter from soils or even the glues used to assemble fossils, the researchers treat the bone with chemicals that tear collagen's triple helices into single chains to release the trapped contaminants. A molecular sieve then filters out contaminating carbon molecules, leaving behind pure collagen. The colour of the final product is a good indicator of its quality, Higham says, holding up a glass bottle containing a white, fluffy, grape-seed-size fleck that resembles cotton wool.

The Red Lady and remains from other sites in Britain were the first that his lab examined. He has since expanded his search across continental Europe, and in 2007 his team won a £350,000 grant from Britain's Natural Environment Research Council in Swindon to re-date three dozen archaeological sites. The number eventually ballooned to 65.

## OLDER AND OLDER

Like the Red Lady, bones from many sites are turning out to be millennia older than previously thought. Before Higham's work, the oldest human bones in Europe were from the Peștera cu Oase cave in southwestern Romania, dated to around 40,000 years old. Higham and his colleagues have now begun to find older examples. In November 2011, they announced that they had dated what would become the oldest human fossil in Britain<sup>5</sup>. A fragment of jaw bone had been discovered in 1927 in Kent's Cavern, a coastal cave in Devon, and had been dated in the late 1980s to about 35,000 years old<sup>6</sup>.

Higham's team assert that the jaw is more than 41,000 years old<sup>5</sup>, on the basis of dates of animal bones excavated above and below the jaw. (The team was unable to date the jaw itself.) Work by Katerina Douka, an archaeological scientist at Oxford (and Higham's partner), published on the same day<sup>7</sup> dated molars from Cavallo Cave in Italy's heel at between 43,000 and 45,000 years old, making them the earliest modern human fossils in Europe, although not everyone agrees that they are human.

"We're starting to build up a picture that modern humans were getting into Europe much earlier than we thought," says Chris Stringer, a palaeoanthropologist at London's Natural History Museum and co-author of the Kent's Cavern paper<sup>5</sup>.

## "IT JUST BREAKS YOUR HEART TO SEE WHAT PEOPLE HAVE DATED BEFORE."

These early incursions may have put humans in direct contact with Neanderthals who had lived there for millennia. "Getting people up to Kent's Cavern near Plymouth, that's a hell of a thing at 40,000 years ago," says Richard Klein, an archaeologist at Stanford University in California. He doubts that they coexisted for long: "It's hard to imagine they were playing games with Neanderthals when they went up there. They must have replaced them very quickly."

Higham says that his dates tell a more nuanced story. He likens Palaeolithic Europe to a giant chess board, with established Neanderthals facing a series of intrusions by modern humans. In places, the two may have lived alongside each other for thousands of years, opening up the possibility of cultural and even sexual exchanges.

Comparisons of modern human genomes with Neanderthals' suggest that some interbreeding occurred (see page 33). But because Asians and Europeans have identical levels of Neanderthal DNA, geneticists presume that they are seeing the result of trysts that occurred before modern humans moved to Europe. Higham's work could help to pin down when and where humans and Neanderthals were most likely to have interbred.

## INVADING EUROPE

Better purification techniques in radiocarbon dating have pushed back the arrival time of the earliest humans in Europe, and could reveal how they crossed paths with Neanderthals.

### KENT'S CAVERN, UK

A fragment of jaw bone dated to 35,000 years ago in 1989 is now Britain's earliest human fossil, at 41,000–44,000 years old.

### VINDIJA, CROATIA

Neanderthal samples were dated to around 33,000 years old. Re-dating pushed the date back further, and suggested that Neanderthals were gone by 40,000 years ago.

### GROTTA DEL CAVALLO, ITALY

Teeth, originally thought to be from Neanderthals, actually provide earliest evidence of modern humans in Europe, 43,000–45,000 years ago.

### EL SIDRON, SPAIN

Neanderthal bones have been dated to an implausible 10,000 years old. Re-dating puts their age closer to 50,000 years old.

### GROTTE DU RENNE, FRANCE

Decorative items in the same layer as Neanderthal remains have been linked to symbolic behaviour. Recent dating questions that interpretation.

go.nature.com/nmz8gd

João Zilhão, a palaeoanthropologist at the University of Barcelona in Spain, has emerged as Higham's staunchest critic. Last year, Zilhão and his colleagues pointed out that the artefacts in the Châtelperronian layer seemed to be in the right place and questioned whether Higham's team had managed to fully decontaminate the bone samples<sup>12</sup>. "How come the bones move and the stone tools do not? It's impossible," he asks. Higham struck back<sup>13</sup>, and Zilhão is now drafting another response. "This could go on forever and I've got no more time to spend on it," says Higham.

Both say that their dispute is purely academic. They continue to work together on other material, and are open to collaboration on the Grotte du Renne controversy. "He's pretty easy to work with," Zilhão says of Higham. "He speaks his mind, but so do I."

Stringer says that the understanding of palaeolithic history is in flux. The dates that Higham and others are now generating may settle some long-standing debates, but they are also generating new questions. "Maybe you've got a muddying of the waters before they clarify and settle out," Stringer says.

## A CINEMATIC VISION

This summer, Higham will trek to the Denisova cave in southern Siberia's Altai Mountains, to try to make sense of its convoluted history. When Soviet scientists found the cave in the 1970s, they discovered Neanderthal tools and human remains there. But in 2010, DNA sequencing of a finger bone extracted from the cave pointed to the existence of a hitherto

unknown population of archaic humans, called Denisovans<sup>14</sup>, who lived in the cave sometime between 30,000 and 48,000 years ago<sup>15</sup>. Higham thinks that his team can narrow down that range and perhaps determine whether Denisovans lived in the region with humans and Neanderthals.

Higham's grand vision is to develop a fuller, almost cinematic version of early human migrations. "We want to create this huge map that will allow us to try to look at the movement of people, the movement of objects, the development of new ideas. The big archaeological questions, really." His team has already begun to play around with software capable of building such a map of Europe, some of which incorporates data from a stack of manuscripts on his desk that he hopes will be published over the next year and a half.

But if this film is to be more historical documentary than a period drama, it requires the sort of chronologies that Higham and his team are generating. "You have to know the dates," he says. ■ [SEE EDITORIAL P.6](#)

**Ewen Callaway** writes for Nature from London.

He thinks that Neanderthals probably went extinct gradually. His work re-dating Neanderthal sites in Croatia<sup>8</sup> and the Caucasus<sup>9</sup> suggests that Neanderthals disappeared from these regions by about 40,000 years ago. Other researchers say that the last Neanderthals may have eked out a living in the Iberian peninsula until as recently as 24,000 years ago<sup>10</sup>, although Higham and his former graduate student, Rachel Wood, have unpublished work that questions that timing.

Still, the part of Higham's work that has generated the most debate (or at least the most journal pages), involves the cognitive abilities of Neanderthals. Neanderthals may no longer be written off as knuckle-dragging brutes, but archaeologists disagree over whether Neanderthals were capable of the sort of symbolic representations that underlie language, art and religion.

Shell beads and other ornaments suggest that modern humans made symbolic objects as early as 100,000 years ago in Africa, and probably carried those traditions with them into Europe. Evidence that Neanderthals were capable of symbolic thinking comes partly from what is known as the Châtelperronian industry in central and southeastern France, which included ornamental objects such as perforated animal teeth, shell beads and ivory pendants. Neanderthal bones found alongside such artefacts at the Grotte du Renne in central France made the site "the flagship for the idea that Neanderthals had symbolic behaviour", says Stringer.

Higham, however, questions how good that evidence is<sup>11</sup>. His team dated animal bones, antlers and teeth from various layers of the cave. The dates for those in the Châtelperronian layers were all over the place, from 49,000 to 21,000 years old. Higham thinks that bones and artefacts from different periods have become jumbled, through a combination of geological tumult, excavation errors and shoddy record-keeping. He therefore doesn't think the Châtelperronian objects should be used to support symbolic thinking for Neanderthals.

- Oakley, K. P. *Antiquity* **42**, 306–307 (1968).
- Hedges, R. E. M., Housley, R. A., Law, I. A. & Bronk Ramsey, C. *Archaeometry* **31**, 207–234 (1989).
- Nyerup, R. *Oversyn over foedrelandets mindesmaerker fra oldtiden* (1806).
- Anderson, E. C. *et al. Science* **105**, 576–577 (1947).
- Higham, T. *et al. Nature* **479**, 521–524 (2011).
- Hedges, R. *et al. Archaeometry* **31**, 207–234 (1989).
- Benazzi, S. *et al. Nature* **479**, 525–528 (2011).
- Higham, T. *et al. Proc. Natl Acad. Sci. USA* **103**, 553–557 (2006).
- Pinhasi, R. *et al. Proc. Natl Acad. Sci. USA* **108**, 8611–8616 (2011).
- Finlayson, C. *et al. Nature* **443**, 850–853 (2006).
- Higham, T. *et al. Proc. Natl Acad. Sci. USA* **107**, 20234–20239 (2010).
- Caron, F. *et al. PLoS ONE* **6**, e21545 (2011).
- Higham, T. *et al. Before Farming* **2011/2012**, 1 (2011).
- Krause, J. *et al. Nature* **464**, 894–897 (2010).
- Derevianko, A. *et al. Archaeol. Ethnol. Anthropol. Eurasia* **34**, 13–25 (2008).





# COMING TO AMERICA

BY ANDREW CURRY

*For decades, scientists thought that the Clovis hunters were the first to cross the Arctic to America. They were wrong — and now they need a better theory.*

T

he mastodon was old, its teeth worn to nubs. It was perfect prey for a band of hunters, wielding spears tipped with needle-sharp points made from bone. Sensing an easy target, they closed in for the kill.

Almost 14,000 years later, there is no way to tell how many hits it took to bring the beast to the ground near the coast of present-day Washington state. But at least one struck home, plunging through hide, fat and flesh to lodge in the mastodon's rib. The hunter who thrust the spear on that long-ago day didn't just bring down the mastodon; he also helped to kill off the reigning theory of how people got to the Americas.

For most of the past 50 years, archaeologists thought they knew how humans arrived in the New World. The story starts around the end of the last ice age, when sea levels were lower and big-game hunters living in eastern Siberia followed their prey across the Bering land bridge and into Alaska. As the ice caps in Canada receded and opened up a path southward, the colonists swept across the vast unpopulated continent.

Archaeologists called these presumed pioneers the Clovis culture, after distinctive stone tools that were found at sites near Clovis, New Mexico, in the 1920s and 1930s.

As caches of Clovis tools were uncovered across North America over subsequent decades, nearly all archaeologists signed on to the idea that the Clovis people were the first Americans. Any evidence of humans in the New World before the Clovis time was dismissed, sometimes harshly. That was the case with the Washington-state mastodon kill, which was first described around 30 years ago<sup>1</sup> but then largely ignored.

Intense criticism also rained down on competing theories of how people arrived, such as the idea that early Americans might have skirted the coastline in boats, avoiding the Bering land bridge entirely. "I was once warned not to write about coastal migration in my dissertation. My adviser said I would ruin my career," says Jon Erlandson, an archaeologist at the University of Oregon in Eugene.

But findings over the past few years — and a re-examination of old ones, such as the mastodon rib — have shown conclusively that humans reached the Americas well before the Clovis people. That has sparked a surge of interest in the field, and opened it up to fresh ideas and approaches. Geneticists and archaeologists are collaborating to piece together who came first, when they arrived, whether they travelled by boat or by foot and how they fanned out across the New World.

To test their ideas, some researchers are examining new archaeological sites and reopening old ones. Others are sifting through the DNA of modern people and unearthing the remains of those buried millennia ago in search of



## PEOPLING THE PLANET

Interactive map of migrations:  
[go.nature.com/nmz8gd](http://go.nature.com/nmz8gd)



genetic clues. “There’s a powerful meshing of the archaeology we’re pulling out of the ground with genetic evidence,” says Michael Waters, a geographer at Texas A&M University in College Station.

Like those original Americans, researchers are exploring new frontiers, moving into fresh intellectual territory after a long period of stasis. “Clovis has been king for 50 years, and now we have to reimagine what the peopling of the New World looked like,” Erlandson says. “If it wasn’t Clovis, what was it?”

## OVERTHROWING KING CLOVIS

It took a chance finding halfway around the world to set this reappraisal in motion. In the late 1970s, Tom Dillehay, an archaeologist at Vanderbilt University in Nashville, Tennessee, uncovered the remains of a large campsite in southern Chile, close to the tip of South America (see ‘Routes to a new world’). Radiocarbon dating of wood and other organic remains suggested that the site was around 14,600 years old, implying that humans made it from Alaska to Chile more than 1,000 years before the oldest known Clovis tools<sup>2</sup>. But because the remote site was so hard for most researchers to examine, it would take nearly 20 years for Dillehay to convince his colleagues.

The case for pre-Clovis Americans has now gained more support, including from analyses of ancient DNA. One of the first bits of genetic evidence came from preserved faeces, or coprolites, that had been discovered in a cave in south-central Oregon by Dennis Jenkins, an archaeologist at the University of Oregon. Radiocarbon dating showed that the coprolites are between 14,300 and 14,000 years old, and DNA analysis confirmed that they are from humans<sup>3</sup>. The recovered DNA even shared genetic mutations with modern Native Americans.

Since the coprolite evidence emerged, in 2008, ancient DNA has also been used to reconstruct that long-ago mastodon hunt. Radiocarbon studies in the 1970s had suggested that the mastodon pre-dated the Clovis people, but some researchers explained that away by arguing that the animal had died in an accident. However, DNA studies last year<sup>4</sup> showed that a fragment of bone embedded in the mastodon’s rib had come from another mastodon — strong evidence that it was a spear point made by humans and not a shard that had chipped off a nearby bone in a fall.

The case against Clovis got another major boost last year, when an excavation in Texas unearthed stone tools that pre-dated Clovis-style artefacts by more than a millennium<sup>5</sup>. “We found a solid site with good context, good artefacts and solid dating,” says Waters.

This slow avalanche of findings has all but buried the Clovis model — the problem now is what to replace it with. The abundant Clovis artefacts and sites discovered over the past century have set a high bar. Telling the story of the first Americans means coming up with a plausible explanation and definitive evidence to support it — a combination that researchers are struggling to achieve.

One idea they are exploring is that a small group of big-game hunters made it into the Western Hemisphere over land — but significantly earlier than previously thought. Another, more popular, theory argues that humans used boats to navigate along the coast of Siberia and across to the Americas.

There is also a controversial variant of the coastal migration model, put forward by archaeologists Dennis Stanford at the Smithsonian Institution in Washington DC and

Bruce Bradley at the University of Exeter, UK. Called the Solutrean hypothesis, it suggests that coastal migration from Asia could have been supplemented by parallel migrations across the Atlantic, bringing stone-tool technologies from present-day Spain and southern Europe to eastern North America.

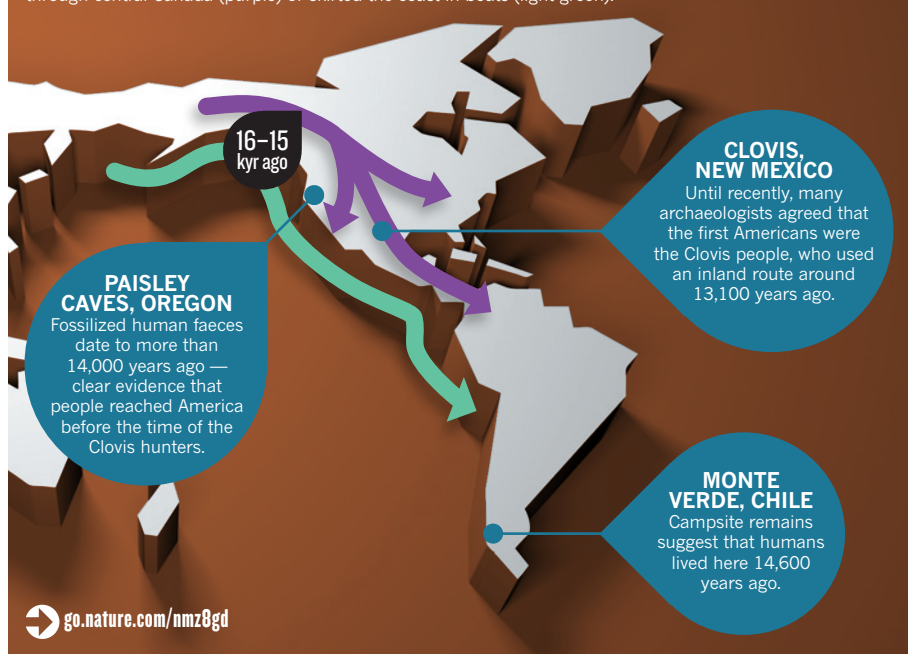
DNA studies argue strongly against this hypothesis, and it gets little support from researchers. But some are hesitant to reject the idea outright, recognizing that the community was once before too conservative. “That’s what happened with the Clovis paradigm,” says Dillehay.

To move the field forward, researchers are using as many types of data as possible. Some key clues have emerged from studies of population genetics, in which researchers tallied the number of differences between the genomes of modern Native Americans and those of people living in Asia today. They then used estimates of DNA mutation rates as a molecular clock to time how long the diversity took to develop. That provides an estimate for when people split from ancient Asian populations and migrated to the Americas.

Judging from the limited genetic diversity of modern Native Americans, Ripan Malhi, a geneticist at the University of Illinois at Urbana-Champaign, and others have argued that the founding population was small, perhaps just a few thousand hardy settlers. In a study of mitochondrial DNA from modern Native American and Asian populations, Malhi and his colleagues also found hints that the first American colonists paused on their way out of Asia<sup>6</sup>, waiting out the peak of the last ice age on the exposed Bering land bridge for perhaps 5,000 years — long enough to become genetically distinct from other Asian populations. When the glaciers blocking their path into North America began to melt around 16,500 years ago, the Beringians made their way south over land or sea, passing those genetic differences on to their descendants in America.

## ROUTES TO A NEW WORLD

Using archaeological remains and genetic sequences, researchers are trying to reconstruct how and when modern humans reached the Americas. The debate centres on whether people walked down through central Canada (purple) or skirted the coast in boats (light green).



**“NOW WE HAVE TO REIMAGINE WHAT THE  
PEOPLING OF THE NEW WORLD LOOKED  
LIKE. IF IT WASN’T CLOVIS, WHAT WAS IT?”**

Other researchers say that there is a major problem with relying on population genetics to answer questions about the peopling of the Americas. At least 80% of the New World's population was wiped out by disease, conflict or starvation after Europeans first arrived some five centuries ago. And the genes of many Native Americans today carry European and African markers, which confounds efforts to piece together the migration story. "If we look pre-contact, we're going to find a lot more indigenous diversity," says Malhi.

That means going back in time, by studying ancient genomes. "You're going to see a lot of ancient-DNA studies coming out, and that's going to tell a powerful story about the first Americans," says Waters.

The chances of finding well-preserved bones from the first Americans are slim, but valuable information can be pulled from DNA samples that fall in between then and now, argues Eske Willerslev, who studies ancient DNA at the University of Copenhagen. Willerslev and his colleague Thomas Gilbert proved that point in 2010, when they extracted the first complete ancient-human genome from a 4,000-year-old hank of hair found in Greenland that had languished for decades in a museum storeroom in Copenhagen. The DNA helped to show that there had been multiple waves of migration into Greenland, and that modern Greenlanders arrived more recently<sup>7</sup>. Now, Willerslev's lab is trying to extract similar information about population movements from ancient-human remains from sites all over the Americas.

## JOINING FORCES

When paired with sequences from modern populations, ancient DNA can help to refine the calculations made by population geneticists and test the claims made by archaeologists. In 2008, Brian Kemp, now at Washington State University in Pullman, extracted mitochondrial DNA from a 10,300-year-old tooth found in On Your Knees Cave in Alaska. When he compared the DNA sequences with those from modern Native Americans, he found that the mutation rate was faster than previously thought<sup>8</sup>. The results, he says, effectively rule out the possibility that humans came to North America as early as 40,000 years ago — a date based on equivocal evidence from archaeological sites in the eastern United States. The finding also argues against the idea that people used boats before the thaw to go around the glaciers and come down the coast. Instead, the DNA evidence supports the consensus that people didn't migrate into the Americas — whether by boat or over land — until the end of the last glacial maximum, 16,500 years ago at most.

The DNA told researchers a few more things. The ancient man who died in that Alaskan cave had mitochondrial DNA most closely related to Native American groups living today along the west coast of North America. "Most of the people who descended from that type are still living near the coast," Kemp says. So the first wave of migrants probably came down the coast and then spread east from there, developing tiny variations in their DNA as they went, Kemp says.

Dennis O'Rourke, a geneticist at the University of Utah in Salt Lake City, is using similar comparisons to fill in the map of ancient migrations in the New World. In the past ten years, dozens of similar studies have established a clear trend — comparisons of DNA from modern people with ancient DNA have shown that the geographic distribution of genetic groups in the Americas has been stable for millennia. "The patterns must have been established more than 4,000 years ago," he says. That helps to constrain the timing of when people spread across

the continent and when they stopped migrating, he says.

In Point Barrow, Alaska, O'Rourke recently began studying human remains from a cliff-top cemetery threatened by coastal erosion, where people have been buried for the past 1,000 years. By comparing the samples from ancient Alaskans to populations from Greenland, eastern Canada and elsewhere, O'Rourke hopes to learn more about the colonization of the Arctic, an environment similar to what the first Americans would have encountered towards the end of the last ice age.

O'Rourke's collaborators are also collecting DNA samples from Inupiat people in northern Alaska. By matching up the modern and ancient DNA sequences from that region, they hope to refine the genetic clock and improve estimates for when people arrived in the Americas. Similar work is going on at a cemetery on Prince Rupert Island off northern British Columbia, where local Tsimshian people are working with archaeologists to gather ancient and modern DNA evidence.

While geneticists open up intellectual frontiers, archaeologists are searching for ways to test the migration theories in the field. Direct evidence for coastal migration will be hard to come by, because a rise

in the sea level since the end of the last ice age has flooded the ancient coastlines. But researchers are turning up indirect evidence in many locations. Last year, for example, Erlandson demonstrated that humans lived on California's Channel Islands as far back as 12,200 years ago<sup>9</sup>, which shows that they must have mastered the use of boats before that time.

And at the Monte Verde site in Chile, researchers have found evidence that the ancient occupants were fans of seafood<sup>10</sup>. "Monte Verde has ten different species of seaweed at the site," Dillehay says. "Somebody was intimately familiar with seaweeds and the microhabitats where they could be found." That lends support to



Radiocarbon dating of ancient faeces found in Oregon shows that humans were in North America as early as 14,300 years ago.

the idea that the earliest Americans were seafarers, he says.

Dillehay's recent findings, which came 30 years after the first excavations at Monte Verde, show that previously studied sites can become potential gold mines, says Waters. Because so many sites were either dismissed or forgotten during the 'Clovis-first' era, Waters says that "the field can really be pushed forward by going back and taking a look at sites that were put up on a shelf". He is already planning to reopen sites in Tennessee and Florida, where evidence of pre-Clovis mammoth hunting was uncovered in the 1980s and 1990s.

Geneticists and archaeologists agree that the death of the Clovis theory has injected the field with excitement and suspense. "There's a sense that there was something before Clovis," says Jenkins, whose coprolite study shook the field four years ago. "But what it was and how it led to the patterns that we see in North and South America — that's a whole new ball game." ■ [SEE EDITORIAL P.6](#)

Andrew Curry is a freelance writer in Berlin.

1. Gustafson, C. E., Gilbow, D. & Daugherty, R. *Can. J. Archaeol.* **3**, 157–164 (1979).
2. Dillehay, T. D. *Monte Verde: A Late Pleistocene Settlement in Chile* Vol. 1 (Smithsonian Institution Press, 1989).
3. Gilbert, M. T. P. *et al. Science* **320**, 786–789 (2008).
4. Waters, M. R. *Science* **334**, 351–353 (2011).
5. Waters, M. R. *Science* **331**, 1599–1603 (2011).
6. Tamm, E. *et al. PLoS ONE* **2**, e829 (2007).
7. Rasmussen, M. *et al. Nature* **463**, 757–762 (2010).
8. Kemp, B. M. *et al. Am. J. Phys. Anthropol.* **132**, 605–621 (2007).
9. Erlandson, J. M. *et al. Science* **331**, 1181–1185 (2011).
10. Dillehay, T. D. *et al. Science* **320**, 784–786 (2008).



# COMMENT

**NATURAL HISTORY** Edward Lear's forgotten work on ornithology **p.36**



**EARTH SCIENCE** How rocks and life evolved together on our planet **p.39**

**MUSIC** Philip Glass on Einstein and the unpredictability of opera composition **p.40**

**EMPLOYMENT** The skills gained in PhD training make it worth the money **p.41**

ILLUSTRATION BY CHRISTIAN DARKIN



## What makes a modern human

We probably all carry genes from archaic species such as Neanderthals. **Chris Stringer** explains why the DNA we have in common is more important than any differences.

In many ways, what makes a modern human is obvious. Compared with our evolutionary forebears, *Homo sapiens* is characterized by a lightly built skeleton and several novel skull features. But attempts to distinguish the traits of modern humans from those of our ancestors can be fraught with problems.

Decades ago, a colleague and I got into difficulties over an attempt to define (or, as I prefer, diagnose) modern humans using the skeletal morphology that is preserved in fossils<sup>1</sup>. Our attempt was well-intentioned:

we were trying to set up strict criteria, based on cranial measurements, to test whether controversial fossils from Omo Kibish in Ethiopia were within the range of human skeletal variation today — anatomically modern humans.

Our results suggested that one skull was modern, whereas the other was

non-modern (or, in palaeontological terms, archaic). What I did not foresee was that some researchers who were not impressed with our test would reverse it, applying it back onto the skeletal range of all modern humans to claim that our diagnosis wrongly excluded some skulls of recent populations from being modern<sup>2</sup>. This, they suggested, implied that some people today were more 'modern' than others. Although I disputed what I considered to be a misuse of our test, I had to recognize the dangers inherent in this prescriptive ►



**PEOPLING THE PLANET**

Interactive map of migrations:  
[go.nature.com/nmz8gd](http://go.nature.com/nmz8gd)



► approach to modern human variation.

Today, scientists are facing a similar situation. In 2010, DNA evidence showed that after modern humans left Africa about 60,000 years ago, they bred for a short period of time with archaic humans — and, as a result, some populations today have more archaic genes than others<sup>3</sup>. These genes might be expressed in the phenotype, and may require a rethink about how and when regional variation developed in *H. sapiens*. Those with alternative agendas may also try to use these new data to rank modern human populations in terms of supposedly different degrees of modernity. Already I'm reading blogs that speculate about whether some groups are less 'modern' than others, and I fear that such discussions endanger the considerable progress promised by palaeogenetic research.

So to highlight these issues and steer the debate in a positive direction, I would like to assert that the term modern humans, by definition, equally describes all humans living today. Some of us may have more DNA from archaic populations than others, but the great majority of our genes, morphology and behaviour derives from our common African heritage. And what unites us should take precedence over that which distinguishes us from each other.

## OUT OF AFRICA

For the past 25 years, there has been a fierce debate about whether our species evolved from distinct ancestors who were spread across the Old World (multiregional evolution) or whether, as I have advocated, we evolved recently in only one area — Africa (recent African origin). I have argued that the physical traits of modern humans can be divided into two sets of features, which have

originated through different mechanisms and over different time scales<sup>3</sup>. The first set includes features shared by all living humans — such as a high and rounded skull, small brow ridge and a bony chin on the lower jaw<sup>3</sup> — that make it clear a skull belongs to *H. sapiens*, and not an ancient form. These traits had evolved in Africa by 100,000 years ago, and were exported from there in an exodus that began about 60,000 years ago.

The second set of traits belonging to modern humans include the regional or 'racial' features that differentiate human populations, such as facial shape, form of the eyelids, type of hair, skin pigmentation and physique.

I have argued that these regional features were added as people spread in small numbers after the shared modern human template had evolved in Africa. The processes involved were natural selection (Charles Darwin's favoured evolutionary mechanism in his 1859 *On the Origin of Species*), different sexual or cultural preferences (emphasized in Darwin's 1871 volume *The Descent of Man*), a founder effect as remote regions were colonized by small pioneer groups, and drift, as newly separated populations diverged through random processes.

That dispersal event potentially brought early modern humans into the realms of other hominin populations living outside Africa, including the Neanderthals in western Eurasia, and *Homo floresiensis* and the ancient species *Homo erectus* in Indonesia. The first DNA to be successfully recovered from Neanderthal fossils seemed to support the view

**"It is not yet clear whether the archaic DNA many of us carry is tied to any visible traits."**

that the Neanderthals represented a lineage and species that was separate from all recent humans. However, increasingly complete genomic reconstructions from Neanderthal fossils<sup>4</sup> and from a newly characterized group called Denisovans — so far known only from a single cave in southern Siberia<sup>5</sup> — have shown these hominins to be related populations that descended from the earlier species *Homo heidelbergensis*, which also gave rise to *H. sapiens* (see 'A winding path').

Comparative genomic studies have revealed elements of DNA that are unique to each of the three groups (recent humans, Neanderthals and Denisovans), some of which could be expressed in the phenotype, and that may be related to such things as the physiology of the brain, skin, skeleton and even sperm<sup>4,5</sup>. But the biggest surprise for many researchers was the evidence from whole genome scans that modern humans living outside Africa each carry about 2.5% of their DNA from Neanderthals; furthermore, people living today in Australia and New Guinea (Australasians) carry about 5% of Denisovan DNA<sup>4,5</sup> (see 'Patchwork planet').

The most likely explanation is interbreeding, such as when modern humans emerged from Africa into Eurasia about 60,000 years ago and met Neanderthals, and when the ancestors of Australasians met some Denisovans. Scientists are continuing to study these archaic genomes in detail, and are finding clear variation within and between non-African people in the amount and kind of archaic DNA that survives in their genomes today. In short, people living outside Africa carry different quantities, and distinct remnants, of archaic DNA from those interbreeding events. Some African populations may also contain genes from ancient groups as a result of interbreeding, as shown by recent DNA evidence<sup>6</sup> and 15,000-year-old fossils from Central and West Africa, which have some ancestral traits such as a longer, lower braincase<sup>3</sup>.

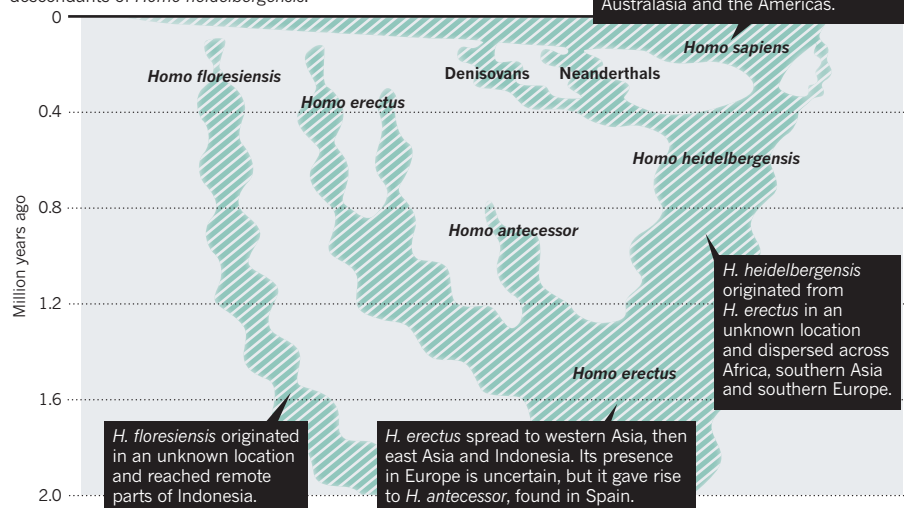
This evidence of interbreeding between groups that palaeontologists call separate species raises two crucial questions. First, given that most of us learned at school that species don't interbreed, should we change the definition we use for a species? Or should we remove the taxonomic separations erected purely from the morphology of fossils, and sink *H. heidelbergensis*, Neanderthals and Denisovans into *H. sapiens*? And second, how does the evidence of interbreeding affect our concept of modern humans, when the genomes of people today apparently contain differing levels of archaic genes?

## LAIED TO REST

In my view, the evidence that *H. sapiens* interbred with archaic humans does not yet require a merging of these close relatives into a single expanded concept of *H. sapiens*.

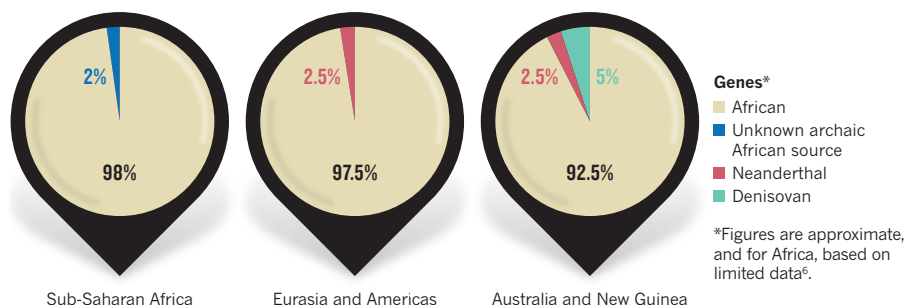
## A WINDING PATH

After early modern humans left Africa around 60,000 years ago (top right), they spread across the globe and interbred with other descendants of *Homo heidelbergensis*.



## PATCHWORK PLANET

Most people's genomes contain remnants of archaic DNA from ancient interbreeding<sup>2-6</sup>.



Doing so would produce a species that had a range of morphological variation several times that found in humans today, or in other existing primate species. These human lineages were distinct enough to build up well-differentiated genotypes and phenotypes (although we know little of the Denisovan phenotype so far) — even the inner ear bones of a Neanderthal are readily distinguishable from those of a modern human. Furthermore, many closely related species of primates undergo limited interbreeding, including among our close ape and monkey relatives<sup>7</sup>. So, for pragmatic reasons, I would retain these species categories while recognizing that this does not imply complete reproductive isolation<sup>3</sup>.

It is not yet clear whether the archaic DNA many of us carry is tied to any visible traits. Are some of the regional physical differences that I argue have largely evolved in the past 60,000 years within dispersing modern humans instead the result of Neanderthal or Denisovan genes? Europeans do not seem to have inherited lighter pigmentation or cold adaptation from their Neanderthal predecessors, but what about their large noses or immune systems<sup>8</sup>? What about the distinctive teeth and malarial resistance of some Australasians — could these be signs of a Denisovan heritage?

More controversially, some of the known differences in coding DNA between Neanderthals and recent humans are associated with brain development and function<sup>4</sup>. Although the archaeological record of Africa suggests that it was the centre of origin for many innovations in modern human behaviour such as complex tools, symbolism and marine exploitation, some have argued that it was only on leaving Africa that modern humans ratcheted up their cognitive skills in response to the environmental challenges of Eurasia<sup>9</sup>. Now, instead, researchers may start to examine the possible phenotypic expression of Neanderthal, Denisovan or African archaic genes for cognition and intelligence.

Terms such as 'archaic' and 'primitive' may be considered objective when used by palaeontologists, but they can be pejorative in common parlance. If researchers want

to continue the progress recently made in studying the origins of modern human variation, they will need to think long and hard about their aims, and the lexicon they use.

One thing should be reiterated: all living humans are members of the extant species *H. sapiens* and, by definition, all must equally be modern humans. The majority of our genes (>90%) derives from our common African heritage, and this should take precedence over the minor amount of DNA that is different — however and whenever it was acquired. It is important that we examine all the factors that lie behind our evolution, including the possible effects of interbreeding on the physiology of modern humans, but we will have learnt nothing in the past 50 years if we let small segments of distinct DNA govern the way we regard regional variation today. ■ **SEE NEWS FEATURE P.23**

**Chris Stringer** is research leader in human origins in the Department of Earth Sciences at the Natural History Museum, London SW7 5BD, UK.

e-mail: c.stringer@nhm.ac.uk

1. Day, M. H. & Stringer, C. B. *L'Anthropologie* **95**, 573–594 (1991).
2. Wolpoff, M. H. Describing Anatomically Modern *Homo sapiens*: A Distinction without a Definable Difference. In *Fossil Man: New Facts, New Ideas. Papers in Honour of Jan Jelinek's Life Anniversary* (eds Novotný V.V. and Mizerová, A.) *Anthropos (Brno)* **23**, 41–53 (1986).
3. Stringer, C. *Lone Survivors: How We Came to be the Only Humans on Earth* (Times Books, 2012).
4. Green, R. E. et al. *Science* **328**, 710–722 (2010).
5. Reich, D. et al. *Nature* **468**, 1053–1060 (2010).
6. Hammer, M., Woerner, A., Mendez, F., Watkins, J. & Wall, J. *Proc. Natl Acad. Sci. USA* **108**, 15123–15128 (2011).
7. Jolly, C. J. *Evol. Anthropol.* **18**, 275–281 (2009).
8. Abi-Rached, L. et al. *Science* **334**, 89–94 (2011).
9. Kohn, M. *The Race Gallery: The Return of Racial Science* (Jonathan Cape, 1995).

## CORRECTION

The article 'A plan for mental illness' (*Nature* **483**, 269; 2012) omitted to provide the full list of authors and the conflicting financial interests declared by some of them. This has been rectified online at [go.nature.com/t1ihrn](http://go.nature.com/t1ihrn).





Lithograph of the scarlet macaw (*Ara macao*) from Lear's parrot monograph, begun when he was 18.

#### NATURAL HISTORY

# The wilder side of Edward Lear

As the poet and artist's bicentenary approaches, **Robert McCracken Peck** celebrates his natural-history legacy.

In 1988, the United Kingdom issued a set of four postage stamps to commemorate the centennial of the death of the beloved artist, poet and travel writer

#### The Natural History of Edward Lear

Houghton Library,  
Harvard University,  
Cambridge,  
Massachusetts.  
Until 18 August 2012.

Edward Lear (1812–88). The stamps featured some of Lear's whimsical ink drawings, including a self-caricature of the bearded artist flying on improbably minuscule wings, and the two boat-bound protagonists of his best-known poem, *The Owl and the Pussycat*.

The affable Lear would have been pleased — and probably astonished — by his country's philatelic attention, but almost certainly disappointed by the choice of images. He considered his illustrated nonsense verse, which ultimately earned him a memorial stone in Westminster Abbey's Poet's Corner, a sideline to his more serious focus: natural-history and landscape painting. Lear so feared that his poetic flights of fancy would undermine his scientific reputation that he hid them behind the pseudonym Derry Down Derry until 1861, long after his nonsense corpus had won a devoted following around the world.

Today, Lear's contributions to science have been mostly forgotten, but early in his life he was a prolific painter of natural-history subjects, earning near-universal praise for his accuracy, originality and style. Lear's greatest scientific contribution was his magnificent *Illustrations of the Family of Psittacidae, or Parrots* (1830–32) — the first monograph to focus on a single avian family — which he began to publish in small batches of individual lithographs when he was just 18 years old. Admired by aviculturists and ornithologists alike, the book's depictions of "species hitherto unfigured" helped to make Lear the artist of choice for many of Britain's leading ornithological publishers in the 1830s and 1840s.

In that golden age of colour-plate books, Lear created some of the most spectacular natural-history illustrations ever published. The original watercolours for these and his other scientific paintings, some of which are on show at the Houghton Library at Harvard University in Cambridge, Massachusetts, have a timeless vitality. Lear met and worked with most of the leading naturalists of his day. He may even have helped John Gould, the chief taxidermist at London Zoo, and his wife Elizabeth — herself a natural-history artist — to create some of the illustrations for Charles Darwin's report on the birds seen during the voyage of HMS *Beagle*.

Yet Lear's later glory masked tough beginnings and ongoing difficulties. The 20th

#### NATURE.COM

For science in the work of Charles Dickens, see: [go.nature.com/79cckns](http://go.nature.com/79cckns)



Lear in his late 20s when, despite deteriorating eyesight, his artistic reputation was thriving.

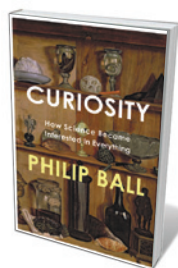
of 21 children, he was born to a prosperous middle-class family in Holloway, now part of north London. But a financial reversal forced the family to disperse when Lear was four years old. He was brought up by his devoted older sister Ann, who gave him lessons that included a rudimentary training in art. As a young boy, Lear earned money “colouring prints, screens, fans” and “making morbid disease drawings for hospitals and certain doctors of physic”, as he noted in the preface to *Nonsense Songs and Stories* (1871).

As a counterweight to this mundane work and to ease his anxiety over the asthma, epilepsy and depression that were to trouble him throughout his life, the youthful Lear created sketchbooks filled with drawings of imaginary birds and animals set in lush, tropical landscapes. A few real-life studies in these hint at the enormous talent he would soon reveal to the scientific world.

Lear's first published illustrations were two vignettes, of lemurs and macaws, in Edward Turner Bennett's *The Gardens and Menagerie of the Zoological Society Delineated* (1830–31). Lear was a natural choice: he had been sketching at London Zoo since it had opened to the public, in 1828.

In June 1830, Lear formally applied for, and received, permission from the Zoological Society's council to draw all the parrots in the society's collection. Over the next two years, he created 42 lithographs for his own folio monograph. Accomplished without institutional, governmental or commercial support, it was an extraordinary achievement, setting new standards for artistic and scientific quality. When renowned ornithologist William Swainson saw the lithograph of a scarlet macaw (*Ara macao*), he wrote to Lear that he considered it equal

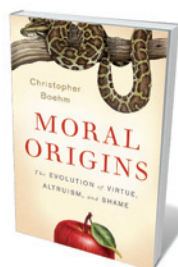
## Books in brief



### Curiosity: How Science Became Interested in Everything

Philip Ball BODLEY HEAD 480 pp. £25 (2012)

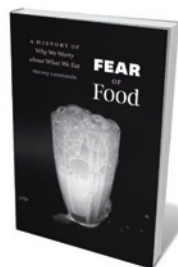
The seventeenth-century shift in Western thought may be labelled the Scientific Revolution but, argues science writer Philip Ball, it was more a sea change in how we view curiosity. That singular trait, which was once condemned, became an engine driving our inexhaustible hunt for knowledge. In this magisterial work, Ball examines the people, events and inventions at the turning point, including “professors of secrets”, cabinets of curiosities, the investigations of Francis Bacon, the plays of William Shakespeare, the rise of polymaths — and the echoes of all this in today's science.



### Moral Origins: The Evolution of Virtue, Altruism, and Shame

Christopher Boehm BASIC BOOKS 432 pp. \$28.99 (2012)

Ethologist and anthropologist Christopher Boehm exposes the roots of generosity and peer disapproval. Eschewing evolutionary game theory, he opts instead for natural selection within the social environment. Boehm posits that selection began with a “crime and punishment” scenario: thugs would have been kicked out of early hunting bands as threats to group survival, and alpha-male hogging of sexual favours would not have been tolerated. Altruism would, argues Boehm, be increasingly valued in partners and allies. Ultimately, such social control might have caused biological change.



### Fear of Food: A History of Why We Worry About What We Eat

Harvey Levenstein UNIVERSITY OF CHICAGO PRESS 232 pp. \$25 (2012)

The United States is a nation gripped by gustatory paranoia, says historian Harvey Levenstein. In this punchy, entertaining account, he reveals how US consumers have suffered for decades from anxiety over the provenance of a pork chop or the fat in fromage fraais. An army of scientists, he says, stoked fear about everything from germs and a lack of vitamins to additives and industrial processing, inadvertently fostering the eating disorder that affects modern US society. Levenstein calls for moderation in all things — including moderation — to regain the joy of eating.



### America's Other Audubon

Joy M. Kiser PRINCETON ARCHITECTURAL PRESS 144 pp. £30 (2012)

Genevieve Jones should have been a towering figure in natural-history illustration. In the 1870s, she resolved to plug the gaps in Jean James Audubon's *The Birds of America* (1838). She learnt lithography, and with Eliza Schulz began the exquisite colour plates for *Illustrations of the Nests and Eggs of Birds of Ohio*. The pictures — sold by subscription — were bought by the likes of US President Theodore Roosevelt and hugely admired. Then Jones died of typhoid. Researcher Joy Kiser tells how Jones's family battled to finish the 68 plates; reproductions, with updated keys, are included.



### The Value of Species

Edward L. McCord YALE UNIVERSITY PRESS 184 pp. £18.99 (2012)

In this meditation on the inherent worth of nature, philosopher and naturalist Edward McCord offers a new twist on the argument for ‘deep conservation’. Each species, he posits, is the astonishing repository of millions of years of natural selection — a volume in Earth's library of life forms. So protecting all creatures is a case of neither need, compassion nor economic value. Instead, their worth is down to their intellectual value, he argues: to the inquisitive mind, each species is incalculably valuable, from bacteria to blue whales.



► in “grace of design, perspective, or anatomical accuracy” to anything by the iconic illustrator John James Audubon; and British naturalist Prideaux John Selby thought the plates “infinitely superior to Audubon’s in softness and the drawings as good”.

The monograph earned Lear election to the Linnean Society of London, but not the financial security that he hoped for. John Gould, who then was just launching his own career as a publisher of ornithological books, saw an opportunity in Lear’s financial distress. He bought unsold copies of the monograph and hired Lear to produce illustrations for several of his own large-format publications. Over the next few years, Lear created 68 powerful plates for Gould’s five-volume masterwork, *The Birds of Europe* (1832–37), and ten illustrations for Gould’s equally beautiful two-volume monograph on toucans (1833–35).

An early admirer was Edward Smith Stanley, later the 13th Earl of Derby and Lear’s most important patron. The earl invited Lear to paint some of the captive birds and mammals living at Knowsley Hall, his sprawling estate near Liverpool. The enormous menagerie contained several thousand animals, including 619 species of birds, along with a natural-history library and some 20,000 mounted and preserved birds and mammals. Lear accepted, and over the next seven years produced more than 100 paintings of the most prized specimens. Seventeen were reproduced in a privately printed book, *Gleanings from the Menagerie and Aviary at Knowsley Hall* (1846), which documented a number of species for the first time.

The exquisite accuracy of Lear’s natural-history paintings was achieved partly by working from living subjects whenever possible. “I am never pleased with a drawing unless I make it from life,” he wrote in 1831. His preparatory studies — stunning watercolours, often heavily annotated regarding colour and form — marvellously mix detail with spontaneity.

That scrupulous accuracy could not hide Lear’s irrepressible bent for the nonsensical. When not painting or socializing with the earl’s guests, he would entertain the children at the hall with a steady flow of whimsy — a way of giving pleasure while



Lear’s ‘hyacinth macaw’ (top) turned out to be an unknown species, now called Lear’s macaw (*Anodorhynchus leari*); bottom, a vitoe (*Nyctipithecus felinus*) from the collection of the 13th Earl of Derby.

cutting across barriers of age and social class. Many of Lear’s famous limericks emerged from these sessions. Many of his nonsense drawings, alphabets and verses drew on his knowledge of natural history: for example, his whimsical sketches of pseudo-plants, such as *Manypeepia upsdownia*, captured the look and sound of real species, both spoofing and celebrating the seriousness of his scientific subjects.

Despite his rapid success, Lear soon tired of scientific illustration. He blamed his loss of interest on failing eyesight, although the detail in some of his later work belies that claim. “My eyes are so sadly worse, that no bird under an ostrich shall I soon be able to see to do,” he wrote at just 24, in the hope of forestalling further illustrative assignments from the demanding Gould.

Lear had long held an interest in landscape painting, and in 1837 Lord Derby agreed to underwrite a two-year study trip to Italy. Lear spent the rest of his life abroad, returning to England only periodically. In 1846, he spent a brief stint as drawing instructor to the young Queen Victoria.

Lear remained a keen observer of nature wherever he was — Italy, Greece, Albania, Egypt, Ceylon (now Sri Lanka) and India, among other destinations — but he never returned to the study of birds and animals, except through his nonsense.

Happily, through one of his early drawings, Lear made an inadvertent scientific discovery of his own. The illustration of the hyacinth macaw (*Anodorhynchus hyacinthinus*) in his parrot monograph was later found to be of a different, previously undescribed species. The ornithologist Charles Lucien Bonaparte, Napoleon’s nephew, named it *Anodorhynchus leari*, or Lear’s macaw, in 1856. This endangered species, which persists in northeastern Brazil, is a living legacy that Lear would undoubtedly have appreciated far more than the nonsense — or the postage stamps. ■

**Robert McCracken Peck** is senior fellow and curator of art and artefacts at the Academy of Natural Sciences of Drexel University in Philadelphia, Pennsylvania, and organized the Houghton Library exhibition.  
e-mail: [peck@ansp.org](mailto:peck@ansp.org)

## EARTH SCIENCE

# How a world came to be

**Birger Schmitz** revels in an account of how life and rock evolved together on Earth.

Without geologists, the Western world would probably still hold to the biblical account of how Earth and life were created. Over the past couple of centuries, thousands have devoted their lives to meticulous studies of rocks — analysing the detailed record of what happened on our planet throughout the ages. In *The Story of Earth*, geologist Robert Hazen summarizes elegantly the modern story of creation.

Drawing on the latest research and influenced by advances in astrobiology, Hazen takes a radical standpoint. He introduces the concept of “mineral evolution” — changes in the composition of Earth’s rocks — to tell the amazing tale of our planet’s intertwined living and non-living spheres.

Hazen’s saga starts 4.6 billion years ago, when the Solar System formed from a gigantic gas cloud, made up mostly of hydrogen with some helium. The Solar System is a relative newcomer to the Universe, which is three times older. Pristine asteroids that condensed when the Solar System formed and later fell to Earth as meteorites tell us about its origin. The young Earth evolved quickly, within a million years of the Solar System’s formation. But 50 million years later, a Mars-sized body dubbed Theia collided with our planet. From the ejecta, the Moon formed.

We have no direct record of what happened on Earth during the planet’s first 800 million years because no rocks remain from this time. But Moon rocks, gathered mostly by NASA’s Apollo missions of the 1970s, reveal that the Moon and Earth were bombarded heavily by huge asteroids. Almost all of the Moon’s craters are from this hellish era; Earth’s surface was probably blasted away during it.

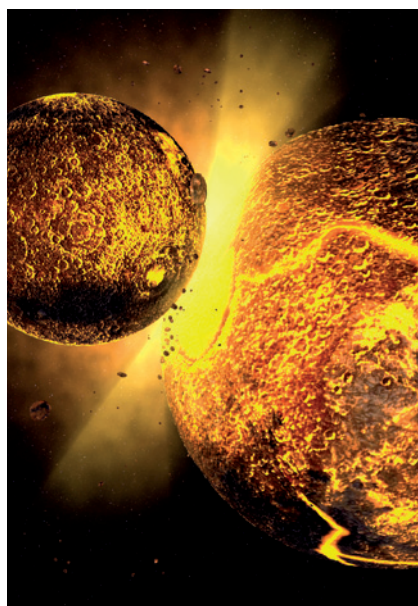
How did life begin? Hazen tells the inside story of an experiment at the University of Chicago, Illinois, in the 1950s. Chemists Harold Urey and Stanley Miller used electric sparks to simulate the effects of lightning on a mixture of boiling water and gases meant to mimic the composition of the early Earth’s atmosphere. A soup of

biomolecules formed, including amino acids and carbohydrates, apparently answering the question of the origin of life. Urey and

Miller’s influence on the field was so strong, Hazen explains, that competing ideas were obstructed for three decades.

The current view is that the building blocks of life can be produced in any environment, given sources of energy and small carbon-bearing molecules. Lightning is not necessary. Amino acids even abound in some meteorites that come to Earth from the apparently sterile Asteroid Belt.

The first single-celled organisms probably originated around 3.8 billion years ago. In one of the most crucial events in Earth’s history, they started to use photosynthesis, and pumped highly reactive oxygen into the atmosphere. About 2.5 billion years ago, the surface of Earth turned from black to red when the basalt crust rusted into soil rich in iron oxide. But it took a much longer time to build up oxygen levels in the oceans.



Earth was shaped by violence, including a collision with Theia that created the Moon.

In a period known as the ‘boring billion’ years, from 1.85 billion to 0.85 billion years ago, the rock record shows no obvious transformative events. The deep ocean was anoxic and inhospitable. But in regions — both at sea and on land — that had been oxygenated, new minerals began to form. Of the roughly 4,500 terrestrial minerals known today, two-thirds could not have formed without oxygen. Beautiful

turquoise, azurite and green malachite exist on a fundament provided by the early single-celled organisms.

From about 0.8 billion to 0.5 billion years ago, Earth experienced its most dynamic transformations. They began with three episodes of near-global glaciations — the Snowball Earth. We do not know why Earth suddenly cooled, but we can be thankful for

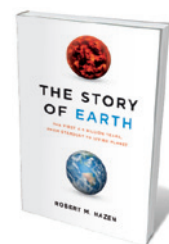
the anomalous behaviour of water. Unlike most compounds, water has a solid phase that is less dense than its liquid phase. If ice did not float, the oceans would have frozen from the bottom up, and would have stayed like that for ever.

During the Snowball Earth, there were no abundant organisms to remove volcanic emissions of carbon dioxide from the atmosphere. A super-greenhouse effect developed and led to rapid melting of the ice. In the wake of Snowball Earth, around 650 million years ago, multicellular organisms evolved. And with the Cambrian Explosion of biodiversity 540 million years ago, all the modern animal phyla arrived. The first organisms with protective shells evolved. The prey–predator game began.

Hazen includes surprises such as that the Earth has an enormous hidden “water ocean”. Until the 1990s, the surface oceans were thought to be the largest water repository, storing about 96% of Earth’s accessible inventory. But studies of interactions between minerals and water under high pressure now tell us that there is 25 times more water stored in Earth’s mantle as in the seas. If it were not for this water there would be no plate tectonics, and no continents. Without continents, no life-sustaining nutrients would have been transported to the oceans by rivers.

Hazen has done his job well. His concise and colourful story of the evolution of our planet and life — how they started from dust, how the evolution of minerals and organisms is intertwined, how the sequence of events depends on countless improbable “if not”s — makes it seem a miracle that we and the rich and diverse nature around us exist. What a wonderful Earth. Let us not destroy what took 4.6 billion years to create. ■

**Birger Schmitz** is professor of geology in the Division of Nuclear Physics at Lund University, Sweden.  
e-mail: [birger.schmitz@nuclear.lu.se](mailto:birger.schmitz@nuclear.lu.se)



**The Story of Earth: The First 4.5 Billion Years, From Stardust to Living Planet**

ROBERT M. HAZEN  
Viking Books: 2012.  
320 pp. \$27.95,  
£17.50





## Q&A Philip Glass

# Relativistic composer

*Einstein on the Beach*, an opera by composer Philip Glass and theatre director Robert Wilson, changed ideas about what opera could do when it was first staged in 1976. As a new production opens at the Barbican Theatre in London, Glass talks about the work's gestation and evolution.

**You have written operas on Galileo Galilei and Johannes Kepler as well as Albert Einstein. Are you drawn to science?**

Science has been a hobby of mine all my life. I took a college course in mathematics, but realized I didn't have the aptitude for science. But it gave me an appreciation for the skills of the scientist. I see scientists as poets, who interpret the world using mathematics. Their visions are personal. Einstein said he would see and understand something, and then figure out the mathematics. Scientists themselves say that there is an aesthetic element to what they do. I see the world as an event, a beautiful unfolding of natural laws. So aesthetics are not confined to artists. Whether you attribute all this to a divine being, as Kepler and Galileo did, or to laws of nature, it doesn't matter.

**How did *Einstein on the Beach* come about?**

Bob Wilson and I wanted to do something together, but we didn't have a name or an idea. We began with the structure — dividing the piece into sections and interludes. And we decided it should be a portrait of a person everyone knew. That way, we didn't need to tell a story — everyone could make up their own. Bob suggested Adolf Hitler, I wanted Mahatma Gandhi, so that didn't work. But

then we came up with Einstein. When I was a boy I had read all about him. We all knew who he was because of his theories of relativity.

Bob began using images from Einstein's work: trains moving, clocks, spaceships.

**But you weren't concerned about a conventional plot?**

No, we didn't have a plot. The libretto was mostly contributed by the performers. And the story is what would be perceived by the spectator — we gave them everything they needed to make it up. This was the 1970s, and we were all influenced by the ideas of choreographer Merce Cunningham, composer John Cage and artist Marcel Duchamp — that the function of interpretation is a creative one, and the creativity comes from the audience.

**What about the structure: five hours, no interval, the audience encouraged to come and go?**

Bob would joke that, if you go to sleep, don't worry — when you wake up, the opera will still be going on. This approach was familiar to me from Kathakali dance-drama in India,

***Einstein on the Beach***

The Barbican Theatre, London.  
4–13 May 2012.

which might start in the evening and finish in the morning.

**Did you draw on the popular myth of Einstein rather than offer straight biography?**

That is exactly right. Einstein was the hanger that we hung the clothes on. We encouraged people to personalize his story — to find in it what was meaningful to them. In the 1940s, there were a lot of popular books explaining relativity. Ordinary people wanted to know what it meant, and Einstein was able to find ways of describing it. Talking about the curvature of space-time creating gravity — that is a complex idea, but if you think about a ball rolling round a hole, it makes it easier to understand.

**What ideas informed the music?**

I had become interested in the rhythmic structure of Indian classical music. It is binary, but based on rhythmic patterns of twos and threes rather than ones and zeros. I've had a long and fruitful association with musician Ravi Shankar, and by the time I was working on *Einstein* I had been developing a whole musical language based on this for ten years. I was also looking for a way to integrate harmonic and rhythmic structure — a kind of unified theory of harmony and rhythm. *Einstein* was an exposition of those ideas.

**Did it work for audiences?**

We first put it on in Europe. We had some income — not enough, but we managed. In the United States, it was staged at the New York Metropolitan Opera and sold out overnight. So we were successful, and totally broke. We had a great time.

**What has changed in the 36 years since then?**

The technology of the theatre has changed dramatically. Lighting instruments are much more advanced, and the movement of scenery can be computer-controlled. Before, you couldn't get enough light in the room to really make the imagery happen. When I walked into the recent rehearsals in Ann Arbor, Michigan, I thought, "Oh my God, this is the image Bob was trying to create." The performances of the musicians and singers are much more refined and accurate than they were back then, and are far closer to what I composed. Bob's skill at moving people has matured too — he has become a master of his language. So the message is there in bold relief.

**And what is the message?**

I'm not going to tell you that! The creative activity of the audience is something I can't do for them. ■

INTERVIEW BY PHILIP BALL

# Correspondence

## PhDs: adapt training for developing world

Paula Stephan condemns the expansion of PhD programmes in the face of “almost no career prospects” (*Nature* **484**, 29–31; 2012). As a mentor of African clinicians undergoing PhD training, I can vouch for the academic rewards that a PhD delivers in developing countries. These include appointments to faculty positions and international grant awards.

African universities are asking for support in developing PhD programmes to train researchers who can address Africa’s myriad problems. To optimize the value of these efforts, they must concentrate on the domains that need most attention and consider the most appropriate type and content of PhD training. Requirements for mentoring, evaluating progress and meeting international standards also need to be set.

**Robert Colebunders** *Institute of Tropical Medicine and University of Antwerp, Belgium.*  
bcoleb@itg.be

## PhDs: acquired skills good for marketplace

Paula Stephan argues that PhD programmes deliver low returns to the taxpayer because career prospects are poor (*Nature* **484**, 29–31; 2012). But at less than 2%, the rate of unemployment for PhDs in the United States is lower than for the general population.

Any highly educated person will be attractive to employers, particularly if that individual has experience in project management, data analysis and effective communication. Jobs in areas such as biomedical research are hard to obtain, but can be found in academia, industry and government. Rewarding careers for PhDs also exist in publishing, intellectual property, science education, conservation and research administration. Thanks to research grants and stipends,

most US graduates will enter the job market with little or no debt.

Much university research is conducted by faculty members and students, rather than by permanent research staff, so new knowledge that could benefit society comes at a relatively low cost to taxpayers.

As with law or medical schools, PhD programmes produce many more graduates than are needed to replace faculty advisers, so earning a PhD does not guarantee a permanent position in either academia or industry. New PhDs face risks, but risk-takers often turn out to be the most innovative investigators.  
**Henrik G. Dohlman** *University of North Carolina, Chapel Hill, USA.*  
henrik\_dohlman@med.unc.edu

## No silver bullets for African soil problems

Conservation agriculture is being used increasingly around the world (H. Buffett *Nature* **484**, 455; 2012). But many of the claimed benefits of minimal or no-tillage farming — such as carbon sequestration and boosting crop yields — are far from proven (K. E. Giller *et al.* *Field Crop Res.* **124**, 468–472; 2011).

Several of these conservation techniques are good agricultural practice: to sow with the first rains, to rotate crops, to fertilize them appropriately and to return crop residues to the soil. But in the absence of herbicides (a reality for many African smallholders), tillage can save labour, allows farmers to plant early and controls weeds. It helps to prevent runoff and erosion if the soil is not protected by mulch, for which smallholder farmers often lack the organic resources.

My view is therefore that an across-the-board recommendation of conservation agriculture is misplaced (*Nature* **483**, 525–527; 2012). Given the enormous diversity of African smallholder farming systems, locally adapted,

best-fit technologies are needed (K. E. Giller *et al.* *Agric. Syst.* **104**, 191–203; 2011).

Farming-systems analysis will identify for whom and where different intensification strategies are most likely to help.  
**Ken E. Giller** *Plant Production Systems, Wageningen University, the Netherlands.*  
ken.giller@wur.nl

## Keep ‘reproducibility’ in context

The call by Glenn Begley and Lee Ellis to raise standards in preclinical cancer research (*Nature* **483**, 531–533; 2012) is an admirable piece of self-critique that can be extended to all who study and treat human disease. Unfortunately, the kernel of their polemic — that reproducing published results can be problematic — was widely misinterpreted by the media.

The authors highlight an investigation in which scientific findings were confirmed in only 6 of 53 cases. This was picked up by news agencies, including Reuters, many of which construed the meaning as

‘findings could not be replicated’ — without mentioning the authors’ nuanced caveats regarding the endpoints tested.

The public cannot be expected to appreciate the implications of, for example, a western blot working in one lab and not in another. Therefore, the meaning of ‘reproducibility’ in a particular context should be properly articulated. This could have prevented the mistaken inference that ‘90% of all science cannot be reproduced’.

In the complex data sets now being generated by genomics and proteomics, the nature of reproducibility is changing; it is increasingly dissimilar from that for more simple measurements. The basic meaning of scientific knowledge is shifting, and the community must learn to deal with it in ways that go beyond simple semantics.

**Brad Picha** *OrthoIndy, Indianapolis, Indiana, USA.*  
**Matthew Thompson Abeon** *Medical, Brecksville, Ohio, USA.*  
**Thomas M. Vondriska** *David Geffen School of Medicine, University of California, Los Angeles, California, USA.*  
tvondriska@mednet.ucla.edu

### EDITORIAL NOTE

In their Comment article ‘Raise standards for preclinical cancer research’, C. Glenn Begley and Lee Ellis (*Nature* **483**, 531–533; 2012) refer to scientists at Amgen who were able to reproduce findings in only 11% of 53 published papers. Several correspondents have asked for details of these studies, which were not provided in the article.

The Amgen scientists approached the papers’ original authors to discuss findings and sometimes borrowed materials to repeat the experiments. In some cases, those authors required them to sign an agreement that they would not disclose their findings about specific papers. Begley and Ellis were therefore not free to identify the irreproducible papers — a fact that the Comment should have mentioned.

*Nature*, like most journals, requires authors of research papers to make their data available on request. In this less formal Comment, we chose not to enforce this requirement so that Begley and Ellis could abide by the legal agreements.

The scientists at Amgen could not have implemented their study had they reserved the right to reveal the outcome for individual papers. The Comment highlights important systemic problems in preclinical cancer research, which we felt appropriate to communicate to our readers, even though the authors could not disclose the studies in question.



## SOLAR SYSTEM

# Focus on ancient bombardment

The latest studies of asteroid impacts on Earth and the Moon beginning about 450 million years after the formation of the Solar System provide insight into the duration, number and size of these events. [SEE LETTERS P.75 & P.78](#)

FRANK T. KYTE

Two papers in this issue present results that have implications for the histories of large-body impacts on the early Earth and Moon. Bottke *et al.*<sup>1</sup> (page 78) propose a significant addition to the widely accepted 'Nice model' for Solar System evolution. Their model greatly increases the number of asteroid impacts expected to have occurred in the Earth–Moon system before 2 billion years ago. Johnson and Melosh<sup>2</sup> (page 75) infer the sizes of ancient impacts by analysing data on ejecta deposits found in ancient sediments. Their analysis uses a computer model<sup>3</sup> that describes how silicate droplets condense from impact-induced vapour clouds\*.

The oldest terrestrial impact craters are only half the 4.56-billion-year (Gyr) age of Earth. Two of the oldest — Sudbury (Canada, 1.85 Gyr) and Vredefort (South Africa, 2.02 Gyr) — are also the largest, with diameters of about 250 and 300 kilometres, respectively. Most impacts older than 2 Gyr are recorded only by spherule beds, deposits of impact ejecta that contain abundant spherules. These are spherical particles quenched from liquid-silicate droplets that condense from clouds of impact-vaporized rocks (Fig. 1). The 0.065-Gyr-old Cretaceous–Palaeogene (K–Pg) boundary layer, for example, which is associated with the mass extinction that included the demise of the dinosaurs, has an iridium-rich global deposit of ejecta (as opposed to regional deposits near the impact) from the 170-kilometre Chicxulub crater in Mexico. This global deposit is a 3-millimetre-thick layer of about 250-micrometre-sized spherules. There are at least 20 known spherule beds, half of which range in age from 2.0 to 3.5 Gyr. Unfortunately, the terrestrial rock record is sparse before 3.5 Gyr ago, and practically non-existent beyond 3.8 Gyr.

The earliest impact history is best preserved on the Moon, but the paucity of samples makes this record difficult to quantify. At least 30 lunar craters have diameters larger than 300 km, the largest being the South Pole–Aitken (2,500 km) and Imbrium (1,145 km) basins. The ages of shock-melted ejecta

\*This News & Views article and the papers under discussion<sup>1,2</sup> were published online on 25 April 2012.



**Figure 1 | Record of an ancient impact.** The image shows a slab cut across the 3.24-billion-year-old spherule bed S3, from the Loop Road region, Barberton Greenstone Belt, South Africa. In this locality, the 25-centimetre-thick bed is a pure deposit of millimetre-scale spherules. Johnson and Melosh<sup>2</sup> estimate that this is the deposit of an asteroid impactor that was 41 to 70 kilometres in diameter. (Photo courtesy of D. R. Lowe and G. R. Byerly.)

from lunar samples collected on the Apollo missions<sup>4</sup>, as well as those of more recently collected lunar meteorites<sup>5</sup>, cluster between 4.1 and 3.8 Gyr. But there is no record of large impacts much older than 4.1 Gyr. This observation led to the hypothesis of a Late Heavy Bombardment (LHB), which implied that, about 450 million years after the formation of the Solar System, the Moon (and Earth) experienced a cataclysmic episode of impacts that ended at about the time that the earliest crustal rocks were preserved on Earth. Because of its larger cross-section, Earth would have about 40 impact basins larger than 1,000 km, and hundreds larger than 300 km.

The Nice model proposes a slow evolution of the orbits of gas-giant planets over several hundred million years that changed when Saturn and Jupiter crossed a 2:1 orbital resonance (at which, for every two orbits of Jupiter there is one orbit of Saturn). When this resonance occurred, about 4.1 Gyr ago, there was a rapid evolution of gas-giant orbits, one result of which was the scattering of asteroids into the inner Solar System, resulting in the LHB. Bottke *et al.*<sup>1</sup> note that the present configuration of the main asteroid belt that lies between Mars and Jupiter, with

an inner edge at 2.1 astronomical units (AU) from the Sun (1 AU is the average distance between Earth and the Sun), is an end product of the reorganization of the planets and the resulting LHB.

Bottke and colleagues reasonably propose<sup>1</sup> that, before about 4.1 Gyr ago, the asteroid belt could have extended in towards the orbit of Mars, to 1.7 AU. When they include this inner 'E-belt' (between 1.7 and 2.1 AU) in the Nice model, they obtain four times as many lunar impact basins caused by objects from the asteroid belt during the LHB. The E-belt asteroids have a greater probability of hitting Earth and the Moon, and higher velocities, than asteroids from the main belt. Importantly, in the E-belt model, many basin-sized impacts would persist on Earth over the succeeding 2 to 3 Gyr. Bottke *et al.* argue that their E-belt model successfully predicts the number of spherule beds and large craters on Earth after 3.5 Gyr ago.

Johnson and Melosh<sup>2</sup> use a model for impact spherule formation which indicates that spherule sizes result from a complex function of the diameter and velocity of impacting asteroids or comets. However, the

**NATURE.COM**  
For more on ancient  
asteroids, see:  
[go.nature.com/ftzclf](http://go.nature.com/ftzclf)

thickness of the global spherule deposit is largely a function of projectile size. The authors apply this model to published data (see Table 1 of the paper<sup>2</sup>) on the size distributions and thicknesses of 14 well-studied spherule beds to estimate probable ranges for projectile sizes and velocities. They find that many of these spherule beds formed from very large projectiles, in the range of 17 to 70 km in diameter (the K–Pg asteroid that formed the Chicxulub crater was 9–14 km). Many of the velocity ranges are rather high for typical asteroid impacts (20 km per second), as much as 21–25 km s<sup>-1</sup>. Perhaps this is consistent with the higher velocities of Bottke and colleagues' E-belt impactors, which are now depleted and no longer significant for the modern Solar System.

For researchers familiar with the spherule beds, these two papers<sup>1,2</sup> help to confirm what has long been suspected — that some of these deposits record impacts far greater than the K–Pg Chicxulub event. At least three of the

Archaean spherule beds (3.2-Gyr-old beds S3 and S4 and the 2.6-Gyr-old Reivilo bed) have higher concentrations of iridium, and thus meteoritic material, than the global K–Pg layer. If many spherule beds are global deposits, then they have 10–100 times more ejecta than the K–Pg boundary layer. Although the new studies provide theoretical support for impact models of the early Earth and Moon, there is still much to be learned, especially for the lunar impact history.

Not all researchers have accepted the LHB hypothesis, and the number and age of the lunar basins is a matter of dispute that can be resolved only with new lunar sample-return missions. There are also unresolved complexities in the spherule-bed record. Three beds from about 3.2 Gyr ago (S2, S3 and S4) all occurred in a brief (about 20-million-year)<sup>6</sup> interval that would be improbable under the Nice model. And the beds' chromium isotopic compositions<sup>6</sup>, although meteoritic, are distinct from those

in three younger (about 2.5-Gyr-old) spherule beds<sup>7</sup>, suggesting that they are derived from separate populations of impactors. These problems do not refute the models in the papers<sup>1,2</sup>, but they do require more data and even better models to be solved. ■

**Frank T. Kyte** is at the Institute of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, California 90095–1567, USA.  
e-mail: kyte@igpp.ucla.edu

1. Bottke, W. F. *et al.* *Nature* **485**, 78–81 (2012).
2. Johnson, B. C. & Melosh, H. J. *Nature* **485**, 75–77 (2012).
3. Johnson, B. C. & Melosh, H. J. *Icarus* **217**, 416–430 (2012).
4. Tera, F., Papanastassiou, D. A. & Wasserburg, G. J. *Earth Planet. Sci. Lett.* **22**, 1–21 (1974).
5. Cohen, B. A., Swindle, T. D. & Kring, D. A. *Science* **290**, 1754–1756 (2000).
6. Kyte, F. T., Shukolyukov, A., Lugmair, G. W., Lowe, D. R. & Byerly, G. R. *Geology* **31**, 283–296 (2003).
7. Simonson, B. M. *et al.* *Precamb. Res.* **175**, 51–76 (2009).

## DRUG DISCOVERY

# Time in a bottle

**A biological clock synchronizes animal behaviour and physiology with Earth's 24-hour rotation. Drugs targeting the clock's 'gears' show promise for treating obesity and other metabolic disorders. SEE ARTICLE P.62 & LETTER P.123**

JOSEPH BASS

**W**hen spring approaches, many countries set their clocks forward by one hour. The following morning, we are reminded of the strong pull exerted by our own internal clock as we are forced to wake up one hour earlier than usual. Such minor disruption of sleep patterns can lead to fatigue and is associated with a rise in the incidence of certain heart disorders<sup>1</sup>. These undesirable effects occur because our internal clock — which tracks the circadian (day–night) cycle — controls not only our sleep patterns but also many physiological processes that anticipate the rhythmic environmental changes tied to the Sun's rising and setting. In this issue, Solt *et al.*<sup>2</sup> (page 62) and Cho *et al.*<sup>3</sup> (page 123) illuminate the molecular mechanisms by which the circadian clock regulates metabolism in mice, and provide evidence to suggest that drugs targeting clock components may offer treatment for disorders such as obesity and diabetes.

In the late 1990s, the finding<sup>4</sup> of biological rhythms in cultured fibroblast cells indicated a broad role for the circadian clock in cell physiology. Subsequently, circadian oscillations were observed<sup>5</sup> in the expression of at least 10% of the genome in mouse tissues. At the molecular level, the circadian clock consists of a feedback loop that involves activator

and repressor proteins, and repeats itself every 24 hours. Activators induce the expression of repressors, whereas repressors inhibit activators' expression.

One of these repressors is REV-ERB- $\alpha$ , a member of the nuclear-receptor family of proteins<sup>6</sup>. In addition to controlling the expression of activators' genes<sup>7</sup>, REV-ERB- $\alpha$  modulates the production of lipids and bile acids in the liver<sup>8,9</sup>, and the formation of fat cells<sup>10</sup>. Whereas the activity of most nuclear receptors is induced on binding to specific steroid hormones, REV-ERB- $\alpha$  has been shown<sup>11,12</sup> to bind instead to haem (an oxygen-binding molecule) and, in turn, to regulate haem synthesis<sup>13</sup>. These findings have increased interest in the development of synthetic compounds that, by binding to REV-ERB- $\alpha$ , could modulate the protein's function.

Notwithstanding these advances, the part played by REV-ERB- $\alpha$  in the circadian clock has remained enigmatic because mice lacking this protein show relatively minor defects in their behavioural rhythms<sup>7</sup>. A possible explanation for this is that a closely related protein, REV-ERB- $\beta$ , can compensate for REV-ERB- $\alpha$  deficiency, as suggested by studies<sup>14</sup> of cultured cells.

To understand the circadian functions of the two REV-ERB proteins, Cho *et al.*<sup>3</sup> identified the genomic regions that these repressors

occupy in the mouse liver. This analysis revealed that both proteins bind to regulatory regions of genes encoding not only numerous core components of the clock but also proteins involved in various metabolic pathways. Therefore, REV-ERBs probably control circadian oscillations through effects beyond modulation of the clock activators. Whether the action of REV-ERBs on clock genes is crucial for the oscillations in cellular activity in other organs — such as the brain — requires additional study.

Mice deficient in REV-ERB- $\alpha$  have been shown<sup>7</sup> to have an increased mortality. So, to carry out experiments with mice lacking both REV-ERB proteins, Cho *et al.* used a genetic-engineering technique known as Cre/lox recombination to generate a mouse strain in which the simultaneous deletion of both genes could be experimentally induced in adulthood. The authors monitored these double-mutant mice running in wheels as a test for circadian dysfunction, and found that the animals' running rhythms had a markedly shortened period length when compared with those of control animals. Moreover, the double mutants displayed an altered response to light.

The researchers compared several metabolic parameters of the double-mutant mice with those of control littermates. The mutant mice had elevated blood levels of triglyceride lipids and of glucose, decreased levels of free fatty acids and a lower respiratory exchange ratio (the relative amount of exhaled carbon dioxide and inhaled oxygen). These metabolic alterations are consistent with an increased generation of energy from fat in the mutant mice.

Cho and colleagues' work<sup>3</sup> provides additional evidence for REV-ERBs as central elements of the circadian clock, and demonstrates that these proteins participate in the control of liver metabolism. To gain further insight



into the functions of REV-ERBs, additional analyses of oxidative metabolism — the process by which cells obtain energy from the oxidation of organic compounds — and exercise tolerance in the mutant mice would be needed. For example, metabolic indicators could be monitored across time and under dynamic conditions, such as during a high-fat diet.

Enter Solt and colleagues<sup>2</sup>. Using a high-throughput screen against the entire family of nuclear hormone receptors in cultured human cells, the authors identified a group of related molecules that selectively activated REV-ERB- $\alpha$  and REV-ERB- $\beta$ . Two of these compounds were suitable for studies in mice, and were investigated further.

The researchers found that the compounds reduced the amplitude of the oscillations in clock-gene expression in cultured cells. And, when injected into mice, the drugs repressed the expression of clock genes. Indeed, the treated mice displayed altered wheel-running rhythms in constant darkness, but not under standard light–dark conditions (12 hours light, 12 hours dark). The cause of reduced drug activity under light–dark conditions on the animals' behaviour requires further investigation but may reflect a direct response to light that bypasses the clock mechanisms — an effect known as 'masking'. The authors carried out further studies in cultured cells that support the idea that the drugs' effects in mice are due to activation of REV-ERBs and not to modulation of other proteins.

Solt *et al.* report that, in addition to the actions on the circadian clock, the compounds protected the animals from certain metabolic disorders associated with obesity and high-fat feeding. The treated mice showed resistance to diet-induced obesity and an increased consumption of oxygen, as well as a reduced food intake during the light period when they are usually sleeping. Moreover, the animals had an altered profile of gene expression in the liver, fat and muscle. In particular, changes in the expression of enzymes involved in the metabolism and transport of fatty acids point towards enhanced oxidative metabolism and reduced lipid storage. The drug treatment also ameliorated metabolic alterations in genetically obese mice that lacked the hormone leptin.

Overall, the results reported by Cho *et al.*<sup>3</sup> and Solt *et al.*<sup>2</sup> re-emphasize the tight coupling of the circadian clock with metabolism, and the special role of REV-ERBs as a nodal point in this relationship. They also suggest that these nuclear receptors may repress the expression of more clock components than previously thought.

Furthermore, the studies raise the possibility of 'putting time in a bottle' — the development of drugs to manipulate biological clocks — for the treatment of metabolic disorders. Admittedly, such an effort entails a chicken-and-egg riddle: any compounds targeting REV-ERBs' activities may affect metabolic parameters

either directly, by modulating the expression of metabolic targets, or indirectly, through effects on the clock. Moreover, as REV-ERB proteins are produced in an oscillatory manner, the actions of any drug would be limited to the window of REV-ERB expression. ■

**Joseph Bass** is in the Division of Endocrinology, Metabolism and Molecular Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA.  
e-mail: j-bass@northwestern.edu

1. Janszky, I. *et al.* *Sleep Med.* **13**, 237–242 (2012).
2. Solt, L. A. *et al.* *Nature* **485**, 62–68 (2012).
3. Cho, H. *et al.* *Nature* **485**, 123–127 (2012).

4. Balsalobre, A., Damiola, F. & Schibler, U. *Cell* **93**, 929–937 (1998).
5. Hughes, M. E. *et al.* *PLoS Genet.* **5**, e1000442 (2009).
6. Lazar, M. A., Hodin, R. A., Darling, D. S. & Chin, W. W. *Mol. Cell. Biol.* **9**, 1128–1136 (1989).
7. Preitner, N. *et al.* *Cell* **110**, 251–260 (2002).
8. Raspé, E. *et al.* *J. Lipid Res.* **43**, 2172–2179 (2002).
9. Le Martelot, G. *et al.* *PLoS Biol.* **7**, e1000181 (2009).
10. Wang, J. & Lazar, M. A. *Mol. Cell. Biol.* **28**, 2213–2220 (2008).
11. Raghuram, S. *et al.* *Nature Struct. Mol. Biol.* **14**, 1207–1213 (2007).
12. Yin, L. *et al.* *Science* **318**, 1786–1789 (2007).
13. Wu, N. *et al.* *Genes Dev.* **23**, 2201–2209 (2009).
14. Liu, A. C. *et al.* *PLoS Genet.* **4**, e1000023 (2008).

The author declares competing financial interests. See [go.nature.com/lwcpbv](http://go.nature.com/lwcpbv) for details.

## SOLID-STATE PHYSICS

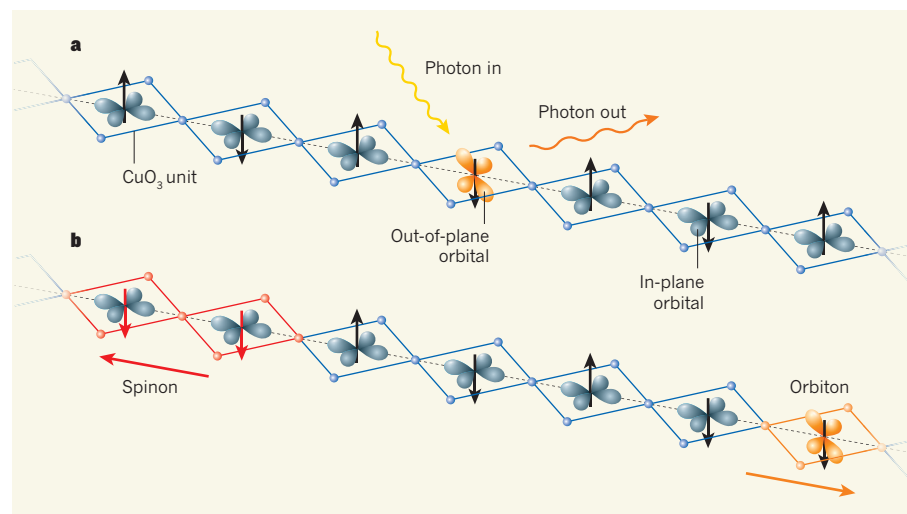
# Electrons do the split

Interacting electrons that are confined to move in a one-dimensional structure do not simply jam together like cars in rush hour. Inelastic X-ray scattering shows that the electrons act as if they split into separate fractional entities. [SEE LETTER P.82](#)

RALPH CLAESSEN

In physics, a phenomenon is often best explained by reducing it to its simplest version. For example, to understand the quantum-mechanical motion of non-interacting electrons in a crystalline solid, the case of a one-dimensional crystal lattice suffices to introduce the concept of electronic band structure, which describes the range of energies that the electrons may have in the solid.

But when it comes to interacting electrons, this approach fails. Coulomb repulsion between any two electrons is much stronger in one-dimensional solids than in their higher-dimensional counterparts, and many-body effects emerge that lead to an apparent fractionalization of the electron<sup>1–5</sup>. With this fractionalization, the electron's spin and charge seem to form separate quasiparticles — spinons and holons, respectively — that move independently of each other and have different velocities. On page 82 of this



**Figure 1 | Electron fractionalization in a one-dimensional structure.** **a**, Schlappa and colleagues' system<sup>6</sup> consists of a chain of copper oxide units ( $\text{CuO}_3$ ). In the ground state, all units contain 3d electrons in the same in-plane orbital, and the direction of the local 3d spins (arrows) alternates between neighbouring units. Shining X-ray photons on the chain excites the electrons in one of the units into an out-of-plane orbital of higher energy. **b**, The authors show that the X-ray irradiation leads to an immediate splitting of the excitation into two separate quasiparticles of different velocities: a spinon, which is associated with a local perturbation in the spin arrangement, and an orbiton, which carries the orbital excitation.

issue, Schlappa *et al.*<sup>6</sup> describe an experiment that takes the idea of electronic fractionalization a step further by revealing that electrons can split into a third form of quasiparticle, the orbiton\*.

Schlappa and colleagues observe orbitons in a copper oxide ( $\text{Sr}_2\text{CuO}_3$ ) containing a one-dimensional chain of copper oxide building blocks ( $\text{CuO}_3$ ), in which the valence electrons reside in the  $3d$  electronic shell of the central copper atoms (Fig. 1). In a free atom there are five  $3d$  orbitals, which are energetically degenerate — that is, they all have the same energy. However, in a  $\text{CuO}_3$  unit, the electrical field of the oxygen ions around each copper atom lifts this degeneracy, generating an intra-shell energy spectrum. In the chain's ground state, the  $3d$  electrons in all of the  $\text{CuO}_3$  units are in the same, lowest-energy  $3d$  state, with their spin orientation alternating between neighbouring units as a result of antiferromagnetic interactions.

In such a system, electrons can be locally excited between the energetically different  $3d$  states, leaving one of the  $\text{CuO}_3$  units in a state known as crystal-field excitation. To achieve this excitation, Schlappa *et al.* shone X-rays on the  $\text{CuO}_3$  chain using a technique called resonant inelastic X-ray scattering. The method is similar to conventional optical Raman scattering, but owing to the short wavelengths of X-rays, it provides information not only about a sample's electronic excitations, but also about the excitations' spatial dynamics.

Traditionally, crystal-field excitations in  $3d$ -metal compounds have been viewed as localized objects that stay fixed to the atomic site at which they were generated. But Schlappa and colleagues<sup>6</sup> provide clear evidence that, in one dimension, such entities break up into two independent, mobile fragments immediately after excitation. These fragments are the spinon, which is a local disturbance of the spin arrangement of the electron ensemble, and the orbiton, a collective response of all  $3d$  electrons in the chain, which carries the  $3d$  intra-shell excitation. The authors find that the separation of the spinon and the orbiton and their ensuing dynamical behaviour are in excellent agreement with theoretical expectations.

The concept of orbitons as collective orbital excitations in  $3d$ -metal compounds, which is similar to that of spin waves in magnetically ordered systems, was introduced more than a decade ago<sup>7</sup>. However, for a long time their propagating character escaped unambiguous experimental verification, because previous work<sup>8</sup> had been directed at materials of higher dimensions than the one-dimensional system studied here. In such systems, intricate coupling of the initial intra-shell  $3d$  excitation to magnetic excitations tends to immobilize the orbitons, preventing their observation

as propagating entities. The spinon–orbiton separation reported by Schlappa *et al.* not only establishes a new aspect of one-dimensional electron fractionalization, but also provides the first clear observation of moving orbitons in a solid.

It should be noted that this observation has become possible only through the enormous advances made in resonant inelastic X-ray scattering during the past few years, notably through pioneering work at the Swiss Light Source facility at the Paul Scherrer Institute in Villigen, Switzerland, where the authors' experiments<sup>6</sup> were conducted. Schlappa and colleagues' results thus highlight the increasingly crucial role of resonant inelastic X-ray scattering in the study of electronic excitations, just as inelastic neutron scattering is the method of choice in the study of magnetic and lattice excitations<sup>8</sup>.

Finally, are there any practical implications for spinon–orbiton separation? As the micro-electronics industry moves towards ever-increasing miniaturization, electronic devices and the conducting

connections between them will eventually reach quantum-mechanical limits. When the diameter of the connections reaches atomic dimensions, the transport of charge in them — and also of spin and heat — will no longer follow the laws for macroscopic (three-dimensional) conductors, and one-dimensional electron fractionalization will become relevant. Whether such fractionalization can be used for new device functionalities in quantum computing or spin electronics remains to be seen. ■

Ralph Claessen is at the *Physikalisches Institut, Universität Würzburg, D-97074 Würzburg, Germany.*  
e-mail: claessen@physik.uni-wuerzburg.de

1. Voit, J. *Rep. Prog. Phys.* **58**, 977–1116 (1995).
2. Kim, C. *et al. Phys. Rev. Lett.* **77**, 4054–4057 (1996).
3. Claessen, R. *et al. Phys. Rev. Lett.* **88**, 096402 (2002).
4. Auslaender, O. M. *et al. Science* **308**, 88–92 (2005).
5. Kim, B. J. *et al. Nature Phys.* **2**, 397–401 (2006).
6. Schlappa, J. *et al. Nature* **485**, 82–85 (2012).
7. Saitoh, E. *et al. Nature* **410**, 180–183 (2001).
8. Ament, L. J. P., van Veenendaal, M., Devereaux, T. P., Hill, J. P. & van den Brink, J. *Rev. Mod. Phys.* **83**, 705–767 (2011).

NATURE.COM  
Find out more about  
quasiparticles at:  
[go.nature.com/fgmbhi](http://go.nature.com/fgmbhi)

#### TRANSLATIONAL MEDICINE

## Double protection for weakened bones

The protein Sema3A both restrains bone degradation and stimulates bone building in mice, suggesting a potential therapy for conditions such as osteoporosis. [SEE ARTICLE P.69](#)

MONE ZAIDI & JAMEEL IQBAL

Every step you take causes microscopic damage to your bones, and your skeleton must therefore be constantly remodelled to maintain its strength. Such remodelling involves the destruction and formation of bone, two processes that are tightly coupled and controlled by signalling molecules such as hormones<sup>1</sup>. A lack of one of these molecules can lead to rapid bone breakdown, as happens when the hormone oestrogen is lost in women after menopause — the resulting osteoporosis accounts for millions of fractures each year. Current treatments for osteoporosis fail to uncouple bone degradation and formation: they stimulate or inhibit both processes at the same time, albeit with different intensities. For these reasons, the results reported by Hayashi and co-workers<sup>2</sup> on page 69 of this issue are truly exciting. The authors find that a protein known as Sema3A facilitates bone regeneration in mice by simultaneously reducing bone

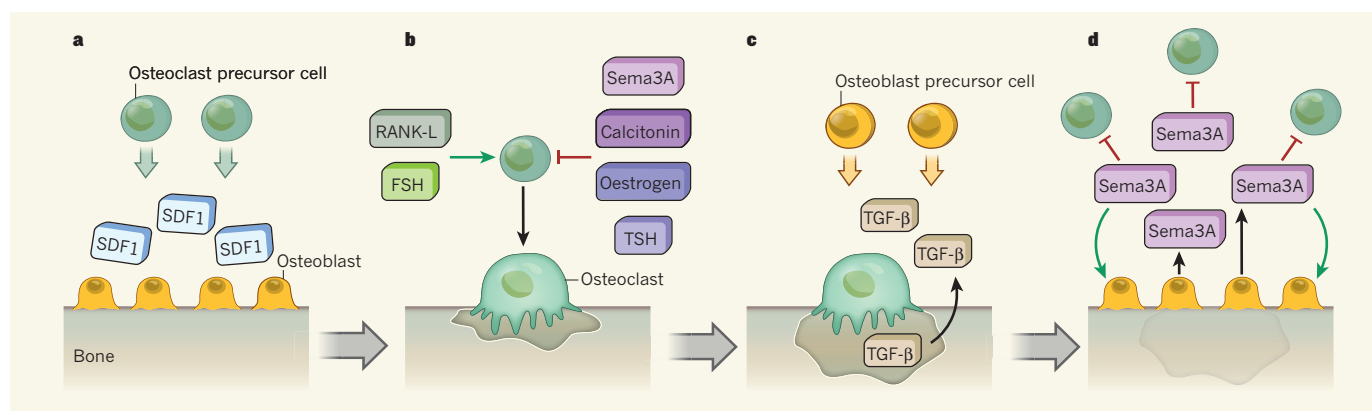
destruction and increasing bone synthesis, and could therefore lead to a new class of dual-action therapeutic agent for osteoporosis.

Bone remodelling is carried out by osteoclasts and osteoblasts — cells that degrade old bone and synthesize new bone, respectively<sup>1</sup>. To search for molecules that could inhibit the degradation process, Hayashi *et al.* analysed the proteins secreted by mouse osteoblasts in culture. The cells had been isolated from mice that lacked the protein osteoprotegerin, a powerful osteoclast inhibitor that would otherwise prevent the detection of less active molecules. Using this technique, the authors identified Sema3A as a potent, direct inhibitor of the formation of osteoclasts from osteoclast precursor cells. As expected, they also found that mice deficient in Sema3A showed features of osteoporosis.

Sema3A is a member of the semaphorin family, a group of proteins involved in the development of the nervous system, organs and blood vessels, as well as in immunity<sup>3,4</sup>. So,

\*This article and the paper under discussion<sup>6</sup> were published online on 18 April 2012.





**Figure 1 | Coupling bone synthesis and degradation.** The activities of osteoblasts and osteoclasts — cells that build and demolish bone tissue, respectively — are tightly balanced. **a**, Osteoclast precursor cells migrate towards specific sites on bone in response to unknown cues. Once at the site, they are further attracted by the protein SDF1. **b**, The balance of several hormones, such as FSH, calcitonin, oestrogen and TSH, and other signalling molecules, such as RANK-L and Sema3A, determines the overall formation of osteoclasts from their precursor cells, and the amount of bone

destruction that ensues. **c**, Osteoclasts solubilize certain molecules, such as TGF $\beta$ , that are usually trapped in the bone matrix; the solubilized molecules then attract osteoblast precursor cells. **d**, Hayashi *et al.*<sup>2</sup> find that newly formed osteoblasts secrete Sema3A to augment and coordinate their own differentiation and bone-building activity. The authors also reveal that Sema3A repels osteoclast precursor cells and inhibits the differentiation of these cells into osteoclasts, thus preventing excessive bone degradation.

how does Sema3A work on bone cells, and can this protein be used in osteoporosis therapy?

The activity of Sema3A on neurons is mediated<sup>4</sup> by its binding to a cell-surface receptor composed of the proteins Nrp1 and PlxnA1. To determine the mechanism by which Sema3A inhibits osteoclast formation, Hayashi *et al.* studied mice expressing an Nrp1 variant that does not bind to Sema3A. The animals showed features of osteoporosis identical to those of the Sema3A-deficient mice, indicating that Sema3A uses Nrp1 as a receptor in bone cells too. Further experiments shed light on how Sema3A exerts its effects on osteoclast precursor cells. Specifically, the authors found that Sema3A binding to Nrp1–PlxnA1 seems to hamper the interaction between PlxnA1 and TREM2–DAP12, a cell-surface receptor that stimulates osteoclast formation and plays a part in the immune response. In the absence of PlxnA1 binding, TREM2–DAP12 cannot activate osteoclast formation.

On the basis of additional work<sup>2</sup>, Hayashi and colleagues suggest that Sema3A not only inhibits the differentiation of precursor cells into osteoclasts, it also repels these precursor cells to prevent excessive bone destruction (Fig. 1). The suggestion is consistent with previous results<sup>4,5</sup> showing that Sema3A inhibits the migration of certain immune cells (thymocytes) and the movement of axonic and dendritic processes of neurons. Interestingly, Sema3A-deficient mice did not seem to have defects in nerve positioning, which suggests that Sema3A's role in this process can be performed by other proteins, whereas its role in bone homeostasis cannot.

Of note, Hayashi *et al.* demonstrate that Sema3A-deficient mice not only have enhanced bone degradation, they also show reduced bone formation. The authors report that the expression of the proteins Rac and

FARP, which form part of the Wnt signalling pathway, is dampened in these mice. As these two proteins usually promote osteoblast formation<sup>1</sup>, their reduced expression probably explains the decreased bone synthesis observed in Sema3A-deficient mice. Interestingly, however, the researchers found that the formation of fat cells (adipocytes) was increased in these mice in comparison to normal littermates. Such a shift from the formation of osteoblasts to that of adipocytes is known to occur<sup>1</sup> in two instances: after activation of PPAR $\gamma$  (a hormone receptor found in the cell nucleus) and with ageing. Hayashi *et al.* note that serum levels of Sema3A decrease with age in mice, but it is unclear whether such a reduction could contribute to the bone loss usually associated with age.

On the therapeutic front, the authors describe how intravenous injection of Sema3A prevents bone loss in a mouse model of menopause. In this model, bone loss is induced by removing the animals' ovaries, the organs that produce oestrogen. Remarkably, the researchers observed<sup>2</sup> that Sema3A injection inhibited bone degradation and promoted bone formation. This mechanism is fascinating, as almost all stimuli or drugs known to decrease bone degradation also reduce bone formation, and vice versa. For example, the injection of parathyroid hormone increases bone synthesis in patients, but bone degradation also rises, minimizing the agent's net effect. Likewise, drugs such as bisphosphonates suppress bone destruction but also reduce bone formation — this forms the basis of clinical concerns underscoring their long-term use in patients. So, the activity of Sema3A is highly unusual because it uncouples the two components of bone remodelling to therapeutic advantage.

If the bone-protection effect is proven in humans, Sema3A could be used not just in the

therapy of osteoporosis but also in that of cancer affecting bone, as suggested previously<sup>6</sup>. Three areas would nevertheless require cautious appraisal. First, although Sema3A — because of its relatively large size — is unlikely to cross the blood–brain barrier and is therefore unlikely to affect nerve positioning in the adult brain, its potential inhibitory effects on the regeneration of peripheral nerves will need evaluation. This is particularly important given that decreased Sema3A expression in injured nerves provides a cue for nerve regeneration<sup>7</sup>.

Second, it has been shown<sup>8</sup> that Sema3A modulates production of vascular endothelial growth factor, a protein with a key role in blood-vessel formation. Although there are no apparent effects of Sema3A on mature vasculature, its potential actions on vessels undergoing rapid remodelling — such as those required for injury repair — would need careful study. Finally, in view of Sema3A's interaction with TREM2–DAP12, noted by Hayashi *et al.*, its potential effects on patients' innate and adaptive immune responses to foreign antigens — for example, during infection, transplantation, vaccination or autoimmune disease — may merit long-term vigilance. ■

**Mone Zaidi and Jameel Iqbal** are at the Mount Sinai Bone Program, Mount Sinai School of Medicine, New York, New York 10029, USA.  
e-mail: mone.zaidi@mssm.edu

1. Zaidi, M. *Nature Med.* **13**, 791–801 (2007).
2. Hayashi, M. *et al. Nature* **485**, 69–74 (2012).
3. Potiron, V. & Roche, J. *Sci. STKE* **285**, pe24 (2005).
4. He, Z. & Tessier-Lavigne, M. *Cell* **90**, 739–751 (1997).
5. Mendes-da-Cruz, D. A. *et al. Exp. Physiol.* <http://dx.doi.org/10.1113/expphysiol.2011.061515> (2012).
6. Chakraborty, G. *et al. PLoS ONE* **7**, e33633 (2012).
7. Pasterkamp, R. J. *et al. Exp. Neurol.* **153**, 313–327 (1998).
8. Maione, F. *et al. J. Clin. Invest.* **119**, 3356–3372 (2009).

## INORGANIC CHEMISTRY

## Two-armed silicon

Compounds of transition metals are often used to activate small molecules for chemical reactions. The discovery of unusual silicon-containing compounds raises the prospect of metal-free activators.

ROBERT WEST

Any chemist will tell you that silicon atoms are tetravalent — that is, they have four electrons available for chemical bonding, and therefore usually form four single bonds to other atoms. But two papers<sup>1,2</sup> published in the *Journal of the American Chemical Society* report some remarkable silicon-containing compounds known as silylenes, in which the silicon atom forms only two single bonds and so is said to be divalent. They are the first stable silylenes that are acyclic (the silicon does not form part of a ring of atoms), and therefore open up a new chapter in the chemistry of divalent silicon.

Among the elements in group 14 of the periodic table (from carbon to lead, all of which have four electrons available for binding), carbon and silicon are the most reluctant to form divalent compounds. The first stable divalent carbon compounds, known as carbenes, were discovered<sup>3</sup> in 1991, and the synthesis of stable divalent silylenes followed a few years later<sup>4–6</sup>. These first, highly reactive silylenes are now regarded as classic compounds, and have become the starting materials for an elaborate branch of silicon chemistry<sup>7–9</sup>.

The classic silylenes are cyclic molecules in which two nitrogen atoms are attached to the silicon. These nitrogens are believed to stabilize the divalent silicon atom, in part by donating electrons to a vacant orbital on the silicon. The cyclic structure also seems to give the molecules stability. Acyclic silylenes that have two nitrogen atoms bonded to divalent silicon have been detected, but these compounds are not stable at room temperature<sup>10,11</sup>. The synthesis of stable acyclic silylenes is therefore a crucial advance because it greatly expands the scope of divalent silicon chemistry.

The compound now reported by Rekken *et al.*<sup>1</sup> has a remarkably simple structure consisting of a silicon atom flanked by two identical, bulky arylsulphur groups (aromatic rings connected to the silicon atom by sulphur atoms; Fig. 1a). The authors prepared it

from a starting material that was structurally identical to the silylene product, but in which two bromine atoms were also attached to the silicon atom. They removed the two bromine atoms using a magnesium-containing reducing agent.

In the second paper<sup>2</sup>, Protchenko and colleagues' silylene (Fig. 1b) has a more complex structure than Rekken and colleagues' compound (for example, it is not symmetrical, and the silicon atom sits between a nitrogen atom and a boron atom), and it required a more complicated synthesis. Protchenko *et al.* also started from a brominated material, but they reacted it with a boron-containing reagent<sup>12</sup> that not only reduces the silicon-bromine bonds in the starting material, but also attaches itself to the silicon atom to yield the final product.

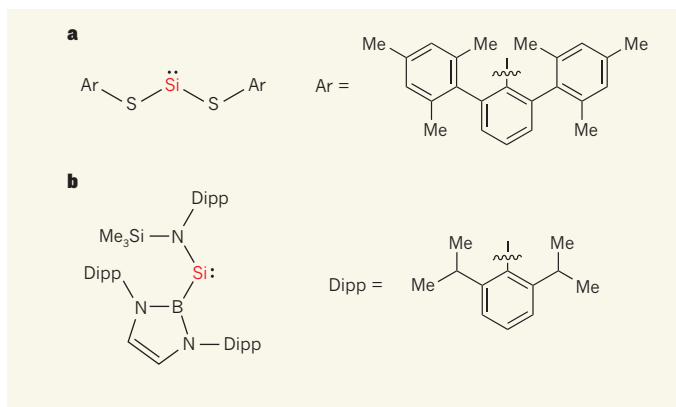
Both Rekken *et al.* and Protchenko and colleagues obtained X-ray crystal structures of their silylenes to provide conclusive proof of what they had made, and also to learn more about the atomic bonding in the compounds — which in turn can help to explain the compounds' reactivities and stabilities. Rekken *et al.* found that the sulphur-silicon-sulphur bond angle in their silylene is 90.52°. The angle at the silicon atom in Protchenko and colleagues' compound is somewhat larger

(109.7°), perhaps implying that their silylene is more reactive than Rekken and colleagues' compound.

Both silylenes exhibit striking thermal stability for divalent silicon compounds: Rekken and colleagues' compound is stable up to 146°C, whereas Protchenko and colleagues' compound is stable up to 130°C. The stability of the new silylenes must be due largely to steric protection of the reactive silicon atoms by the attached groups (in other words, the bulky groups shield the silicon atoms from attack by other molecules). The bulky groups might also stabilize the compounds by donating electrons to the silicon atom.

Further insight into the chemical bonding in the two silylenes was obtained by looking at the signals (the chemical shifts) of silicon-29 nuclei in the nuclear magnetic resonance spectra for the molecules. In both cases, the chemical shifts have strongly positive values (285.5 parts per million for the silicon in Rekken and colleagues' silylene, and 439.7 p.p.m. for that in Protchenko and colleagues' compound). These values fall between the corresponding chemical shift (78 p.p.m.) measured in a classic silylene<sup>4</sup> that is stabilized by electron donation from nitrogen atoms on either side of the silicon, and that of another previously reported silylene<sup>13</sup> that is stabilized only by steric hindrance (567 p.p.m.). The chemical shifts for the new silylenes provide further clear evidence that the compounds are indeed divalent silicon compounds. Moreover, the shifts suggest that Rekken and colleagues' silylene is partially stabilized by electron donation from the sulphur atoms adjacent to the silicon atom, and that Protchenko and colleagues' silylene may also be partially stabilized by electron donation from the atoms on either side of the silicon, but to a lesser extent.

The interest in silylenes lies in their reactivity — particularly the possibility that they can 'activate' small molecules, such as hydrogen, to allow such molecules to take part in potentially useful chemical reactions. Rekken *et al.* report that their silylene does not react with hydrogen, although it does combine with another small molecule (iodomethane; CH<sub>3</sub>I) in a typical silylene reaction. Protchenko *et al.*, however, find that their compound captures hydrogen readily to yield a dihydride product that has two silicon-hydrogen bonds. This is the first example of the reaction of hydrogen with a silylene. What's more, the reaction is a remarkable single-site activation process — one in which both hydrogen atoms become attached to the same atom. Most other activation processes involve two sites. Further study of the chemical



**Figure 1 | Acyclic silylenes.** Two teams report the synthesis of the first thermally stable compounds in which a divalent silicon atom — one that forms two single bonds — does not form part of a ring of atoms. **a**, In Rekken and colleagues' silylene<sup>1</sup>, a silicon atom is flanked by two sulphur atoms. The dots on the silicon atom represent a lone pair of electrons, and the bond crossed by a wavy line in the Ar group indicates the point of attachment of the group to the silylene. Me, methyl group. **b**, In Protchenko and colleagues' silylene<sup>2</sup>, the silicon atom is between a nitrogen and a boron atom. Divalent silicon atoms in **a** and **b** are shown in red.



reactions of these acyclic silylenes will no doubt lead to the synthesis of a variety of interesting chemical compounds. And now that Rekken *et al.* and Protchenko *et al.* have shown the way, we should see a growing family of acyclic, divalent silylenes with as yet unknown structures and properties. ■

**Robert West** is in the Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706-1396, USA.  
e-mail: west@chem.wisc.edu

1. Rekken, B. D., Brown, T. M., Fetting, J. C., Tuononen, H. M. & Power, P. P. *J. Am. Chem. Soc.* <http://dx.doi.org/10.1021/ja301091v> (2012).
2. Protchenko, A. V. *et al.* *J. Am. Chem. Soc.* <http://dx.doi.org/10.1021/ja301042u> (2012).
3. Arduengo, A. J. III, Harlow, R. L. & Kline, M. J. *Am. Chem. Soc.* **113**, 361–363 (1991).
4. Denk, M. *et al.* *J. Am. Chem. Soc.* **116**, 2691–2692 (1994).
5. West, R. & Denk, M. *Pure Appl. Chem.* **68**, 785–788 (1996).
6. Gehrhus, B., Lappert, M. F., Heinicke, J., Boese, R. & Bläser, D. *J. Chem. Soc. Chem. Commun.* 1931–1932 (1995).

7. Hill, N. J. & West, R. J. *Organomet. Chem.* **689**, 4165–4183 (2004).
8. Asay, M., Jones, C. & Driess, M. *Chem. Rev.* **111**, 354–396 (2011).
9. Gehrhus, B. & Lappert, M. F. *J. Organomet. Chem.* **617–618**, 209–223 (2001).
10. Tsutsui, S., Sakamoto, K. & Kira, M. *J. Am. Chem. Soc.* **120**, 9955–9956 (1998).
11. Lee, G.-H., West, R. & Müller, T. J. *Am. Chem. Soc.* **125**, 8114–8115 (2003).
12. Segawa, Y., Yamashita, M. & Nozaki, K. *Science* **314**, 113–115 (2006).
13. Kira, M., Ishida, S., Iwamoto, T. & Kabuto, C. J. *Am. Chem. Soc.* **121**, 9722–9723 (1999).

## CANCER BIOLOGY

# The director's cut

**A genome-wide characterization of active translation of messenger RNA following inhibition of mTOR will transform our view of this signalling protein's regulatory role in cancer. SEE ARTICLE P.55 & LETTER P.109**

ANTONIO GENTILELLA & GEORGE THOMAS

Many cellular responses to external stimuli are comparable to the integrated performances of musicians in an orchestra, as directed by the conductor. By elucidating cellular signalling pathways, scientists can identify the individual components underlying disease, and attempt pharmacological mitigation. These ambitious goals require groundbreaking approaches. One such tool is ribosome profiling — a global approach to assess active translation of cellular messenger RNA that provides an unprecedented picture of the regulation of protein synthesis<sup>1</sup>. In this issue, Hsieh *et al.*<sup>2</sup> (page 55) and Thoreen *et al.*<sup>3</sup> (page 109) use this method to analyse the 'translational landscape' controlled by a kinase protein named mTOR, a key regulator of protein synthesis.

mTOR regulates the translation of many proteins that promote cell-cycle progression<sup>4</sup>, and aberrant mTOR activity is implicated in numerous human diseases, particularly cancer. This means that resolving the complexity of the translational network directed by this kinase could markedly improve our understanding of basic disease processes, and pave the way for more accurate therapeutic strategies.

Previously, mTOR was thought to regulate the translation of a family of mRNAs that have highly structured 5' untranslated regions (5'UTRs) located upstream of the translation start site (TSS) in the mRNA molecule. In a cell in which mTOR signals are absent, members of the 4E-BP protein family suppress the translation of mRNAs by binding the protein-synthesis initiating factor 4E (ref. 4). Activation of mTOR leads to phosphorylation of 4E-BP proteins, which represses their activity, thereby restoring protein synthesis (Fig. 1).

The two new studies<sup>2,3</sup> broaden this view of mTOR activity. The researchers used ribosome profiling in human prostate cancer cells<sup>2</sup> and mouse embryonic fibroblast cells<sup>3</sup> that had been treated with mTOR-inhibitory molecules called PP242 and Torin 1, respectively. The drugs are ATP-site competitive inhibitors that directly inhibit mTOR's ability to bind ATP, a molecule required to functionally modify substrate protein phosphorylation. Because these drugs inhibit mTOR activity more fully than the archetypal mTOR inhibitor rapamycin, which alters the protein's substrate recognition but not kinase activity<sup>5,6</sup>, the researchers were able to thoroughly probe which mRNAs are affected by mTOR inhibition.

Unexpectedly, both groups report that the main class of mRNA molecules affected shortly after drug treatment contain short, unstructured 5'UTRs, in contrast to the highly structured 5'UTRs previously reported<sup>4</sup>. Nevertheless, a parallel study<sup>7</sup> has shown that this latter class of mRNA is negatively affected at later times after mTOR inhibition. Hsieh *et al.* and Thoreen *et al.* identified fewer than 200 affected protein-encoding transcripts, which fall into two groups (Fig. 1). The mRNA sequences of one group are characterized by a tract of pyrimidine RNA bases at the TSS, termed a 5'TOP, which invariably begins with a cytosine base<sup>8</sup>. Hsieh and colleagues identify the second group of mRNAs as containing a pyrimidine rich translational element (PRTE) within the 5'UTR, whereas Thoreen and colleagues describe these molecules as having a 5'TOP-like sequence at the TSS.

The 5'TOP sequence is known to regulate the translation of mRNAs coding for ribosomal proteins and a number of protein-synthesis factors, and it has been shown<sup>9</sup> that the translation of mRNAs containing 5'TOP sequences

is controlled by mTOR. However, these previous studies were performed using rapamycin, which only partially inhibits 4E-BP activity<sup>6</sup>. Importantly, an earlier study<sup>10</sup> found that overexpression of 4E in cells was not sufficient to drive the translation of 5'TOP-mRNAs and that the level of translation of 5'TOP-mRNAs versus other mRNAs showed similar sensitivity to a 4E-inhibitory molecule. These findings suggested that the selective effects of 5'TOP-mRNA translation are independent of 4E and possibly mediated by a regulatory factor that interacts directly with the 5'TOP (ref. 11).

However, the current studies show that neither PP242 nor Torin 1 has an inhibitory effect on translation of 5'TOP mRNAs in cells lacking 4E-BP1/2, despite the fact that these drugs inhibit mTOR more effectively than rapamycin. How can these conundrums be explained? It may be that 4E-BP proteins are required but are not sufficient for 5'TOP translation. Because translation is regulated by proliferative signals and nutrients, it will be important to determine whether deprivation of either suppresses 5'TOP translation in 4E-BP-deficient cells. Moreover, as 5'TOP mRNAs can represent up to 20% of a cell's total mRNA, it will be crucial to determine whether ribosome-profiling techniques such as those used by Hsieh *et al.* and Thoreen *et al.* favour the representation of this class of mRNA over less abundant transcripts with highly structured 5'UTRs.

One of the most intriguing observations in these papers is the identification, by both groups, of new and distinct regulatory motifs (RNA sequence elements found in mRNA molecules) that are regulated by mTOR. Hsieh *et al.* show that the PRTE motif is present in most mRNAs that are sensitive to mTOR inhibition, with most of the molecules also having a 5'TOP, although two small subsets of the mRNAs have only one of the two motifs. In turn, the 5'TOP-like mRNAs identified by Thoreen and colleagues do not seem to contain a cytosine at the TSS, although a complete characterization of the TSS of those mRNAs is still required. In light of these observations, it is tempting to surmise that combinations of regulatory elements within the 5'UTR of these mRNAs may confer differential sensitivity to mTOR signalling.

These findings suggest the existence of a

defined mTOR 'code' that may mediate the translational response governed by different mTOR effectors under different cellular circumstances. A similar translational code, composed of multiple distinct regulatory motifs, has been identified in untranslated regions of mRNAs that are regulated during embryonic development<sup>12</sup>. It is also worth noting that YB1, a tissue-invasion protein encoded by an mRNA that Hsieh *et al.* identify as an mTOR-regulated 5'TOP mRNA, itself acts as a translational regulator, suggesting that there may be hierarchical waves of translational regulation. Intriguingly, control of translation by YB1 correlates with its anticancer activity<sup>13</sup>, raising the question of whether mTOR is also involved

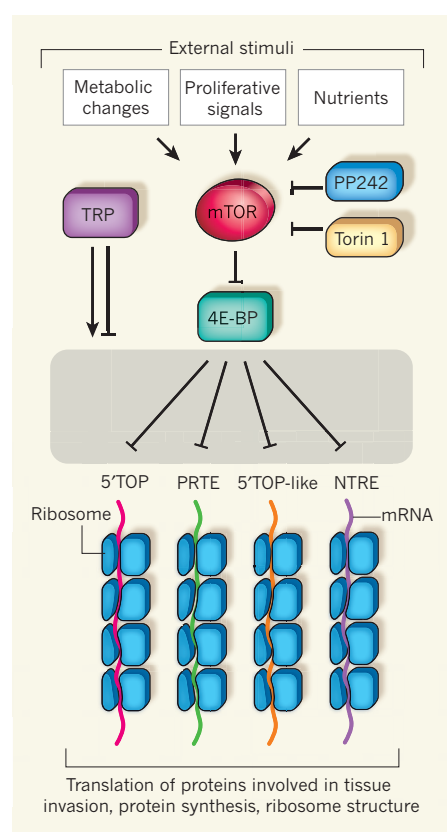
in the YB1-mediated translational response. Experiments designed to silence mTOR activity at different points of signalling pathways and different times during tumour progression will help us to understand the repertoire of translational responses orchestrated by mTOR.

Hsieh and colleagues' study<sup>2</sup> also identified a number of other genes coding for proteins involved in tissue invasion by tumours, including MTA1, CD44 and vimentin, and showed that their translation parallels mTOR activity in prostate tumour cells and in a mouse model of prostate cancer. The researchers also observed this gene-expression pattern in human prostate tumours, and showed that treatment of prostate-tumour-bearing mice with INK128, a drug analogue of PP242, reduces cancer metastasis in the animals. Targeted ATP-site inhibitor drugs have already shown success in the clinic, but unintended effects of off-target kinase inhibition are always a concern. However, in the case of mTOR, the recent finding<sup>6</sup> that RAD001, a rapamycin that binds exclusively to mTOR, acts synergistically with BEZ235, an mTOR ATP-site inhibitor similar to those used in the current studies, to block tumour growth in mice, suggests that such drug combinations may increase the

efficacy of anticancer therapy targeting this pathway. ■

**Antonio Gentilella and George Thomas** are at the Catalan Institute of Oncology, ICO, Bellvitge Biomedical Research Institute, IDIBELL, 08908 Barcelona, Spain. G.T. is also in the Department of Internal Medicine, Metabolic Diseases Institute, University of Cincinnati, USA.  
e-mail: thomasg4@ucmail.uc.edu

- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. *Science* **324**, 218–213 (2009).
- Hsieh, A. C. *et al.* *Nature* **485**, 55–61 (2012).
- Thoreen, C. C. *et al.* *Nature* **485**, 109–113 (2012).
- Dowling, R. J. *et al.* *Science* **328**, 1172–1176 (2010).
- Yip, C. K., Murata, K., Walz, T., Sabatini, D. M. & Kang, S. A. *Mol. Cell* **38**, 768–774 (2010).
- Thomas, H. E. *et al.* *Sci. Transl. Med.* <http://dx.doi.org/10.1126/scitranslmed.3003923> (2012).
- Larsson, O. *et al.* *Proc. Natl Acad. Sci. USA* (in the press).
- Levy, S., Avni, D., Hariharan, N., Perry, R. P. & Meyuhas, O. *Proc. Natl Acad. Sci. USA* **88**, 3319–3323 (1991).
- Jeffries, H. B. *et al.* *EMBO J.* **16**, 3693–3704 (1997).
- Shama, S. *et al.* *Gene Exp.* **4**, 241–252 (1995).
- Ivanov, P., Kedersha, N. & Anderson, P. *Genes Dev.* **25**, 2119–2124 (2011).
- MacNicol, M. C. & MacNicol, A. M. *Mol. Reprod. Dev.* **77**, 662–669 (2010).
- Bader, A. G. & Vogt, P. K. *Oncogene* **27**, 1179–1182 (2008).



**Figure 1 | The translational landscape of mTOR.** The kinase protein mTOR integrates the cellular response to external stimuli by regulating the translation of messenger RNA into proteins that have multiple cellular functions. Hsieh *et al.*<sup>2</sup> and Thoreen *et al.*<sup>3</sup> used a technique called ribosomal profiling — ribosomes are RNA–protein complexes that translate mRNA into protein — to explore, on a genome-wide basis, which mRNAs are regulated by mTOR. By inhibiting mTOR activity using the molecules PP242 or Torin 1, the researchers show that mTOR-regulated mRNAs fall into three types, according to characteristic sequences of RNA bases (5'TOP, PRTE or 5'TOP-like). However, it remains possible that mTOR also regulates the translation of mRNAs containing other new translational regulatory elements (NTRE). In addition, other translational regulatory proteins (TRPs) might influence the translation of these classes of mRNAs.

## EARTH SCIENCE

## Lower mantle may be rich in silica

There have been conflicting claims about the composition of Earth's lower mantle. The latest Brillouin-spectroscopy data suggest that this section of the planet's interior may contain more silica than the upper mantle. [SEE LETTER P.90](#)

IAN JACKSON

In this issue, Murakami *et al.*<sup>1</sup> describe pioneering opto-acoustic measurements of transverse, or shear, elastic waves on polycrystalline specimens of the dominant lower-mantle minerals silicate perovskite, (Mg,Fe)SiO<sub>3</sub>, and ferropericlase, (Mg,Fe)O. Crucially, the researchers performed the measurements under high-pressure and high-temperature conditions comparable to those of the deepest parts of Earth's mantle.

Murakami and colleagues (page 90) find<sup>1</sup> that the shear-wave speeds for both minerals vary more mildly with pressure than had been suggested from extrapolations of data obtained at lower pressures. Consequently, the variation of wave speed with pressure for the perovskite mineral, which displays the higher wave speed, is close to that determined seismologically<sup>2</sup> for the lower mantle, thereby excluding the possibility of a substantial volume fraction of

ferropericlase, which has a much lower wave speed (see Fig. 4b of the paper<sup>1</sup>). Accordingly, the authors conclude that the lower mantle is perovskitic in composition — markedly more silica-rich than the upper mantle. A silica-enriched lower mantle would reconcile the major-element chemical composition of the bulk-silicate Earth with that of carbonaceous chondrite meteorites, which are commonly thought to approximate the composition of the solar nebula from which Earth accreted.

The elegant opto-acoustic technique of Brillouin spectroscopy has, over the past three decades, revolutionized our knowledge of mineral elasticity. In this method, the inelastic scattering of monochromatic light in a crystal, by lattice waves of thermal origin, results in a frequency shift of the scattered radiation that is proportional to the speed of the lattice wave. Systematic variation of the crystal orientation relative to that of the incident light allows determination of the full set of constants



describing the elasticity of a crystal of arbitrarily low structural symmetry. Introduction of this technique to the Earth sciences allowed the single-crystal elasticity of high-pressure silicate minerals to be determined for the first time — from Brillouin spectra measured on 100-micrometre-sized crystals<sup>3</sup>.

Brillouin spectroscopy is also ideally suited to measurements of microcrystals confined at very high pressures within a diamond-anvil high-pressure apparatus<sup>4,5</sup>. Murakami and colleagues' application of the technique to polycrystalline samples of minerals with strongly direction-dependent (anisotropic) elasticity, under conditions of extreme pressure and high temperature, breaks new ground.

Analyses of the elasticity, chemical composition and temperature of the lower mantle date from the classic work of Birch<sup>6</sup>, who boldly suggested that "dense high-pressure modifications of the ferro-magnesian silicates, probably close-packed oxides ... are required to explain the high elasticity of the deeper part of the mantle". This hypothesis was confirmed by subsequent demonstration of the transformation of upper-mantle silicates first to spinel, garnet and related crystal structures<sup>7</sup> at pressures corresponding to mantle depths of 250–500 kilometres, and, ultimately — at depths greater than 660 kilometres — to a mixture of silicate perovskites based on SiO<sub>6</sub> octahedral building blocks, and coexisting ferropericlase<sup>8</sup>. However, the question of whether or not there is also a change in chemical composition between the upper and lower mantle has not, until now, been answered definitively.

Increasingly detailed knowledge of the pressure–temperature conditions required for the stability of the relevant minerals, and of their densities and elastic properties, has provided the basis for numerous analyses of the elasticity, composition and temperature of the lower mantle. The result has been a series of conflicting claims over whether a contrast in chemical composition between the upper and lower mantle is required to match the seismological models. Diverse inferences<sup>9–12</sup> have been drawn concerning the need for lower-mantle enrichment or depletion in silica, iron oxide and other components, and/or for substantial departures from the adiabatic temperature–depth gradient — expected to develop in a fluid subject to thermal convection without exchange of heat between ascending and descending parcels. Such contrasting conclusions are attributable to differences in methodology, trade-offs between chemical composition and temperature, and to residual uncertainties in the thermo-physical database for the high-pressure minerals.

It is gradually becoming accepted that an internally consistent framework<sup>13</sup> should be used for the evaluation and assimilation of thermoelastic data from laboratory experiments and computer models of mineral behaviour, and for extrapolation to higher

pressures and temperatures. Application of such a framework to diverse experimental data for periclase (MgO) highlighted and resolved<sup>14</sup> minor tensions between different data sets. The result was a robust compromise model of the material's thermoelastic behaviour, including a well-constrained value for the pressure derivative — at zero pressure — of the shear modulus, or rigidity, from which the shear-wave speed is calculated.

A markedly lower value (by about 20%) of the zero-pressure pressure derivative was deduced by Murakami and colleagues from ultra-high-pressure Brillouin spectroscopy on polycrystalline MgO (ref. 15). Possible explanations for this discrepancy include a systematic error in pressure calibration in the ultra-high-pressure diamond-anvil experiments and uncertainty concerning the average wave speed that is determined by Brillouin scattering of incident light by the individual crystallites within a polycrystalline specimen.

Given the systematically low values of the pressure derivatives of shear modulus obtained from all of the recent ultra-high-pressure Brillouin spectroscopic measurements, and the consequences for inferred lower-mantle composition, it is vital that the technical issues surrounding pressure calibration and Brillouin scattering from polycrystalline material — ideally of both shear and compressional waves — be resolved. More precise *ab initio* calculations of elastic properties and laboratory measurements independent of an empirical pressure scale<sup>16</sup> may help to explain and to eliminate the discrepancy.

Notwithstanding the impressive experiments of Murakami *et al.*, we are probably still awaiting the final word on the chemical composition and thermal regime of Earth's lower mantle. ■

**Ian Jackson** is at the Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia.  
e-mail: ian.jackson@anu.edu.au

1. Murakami, M., Ohishi, Y., Hirao, N. & Hirose, K. *Nature* **485**, 90–94 (2012).
2. Dzierwinski, A. M. & Anderson, D. L. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
3. Sawamoto, H., Weidner, D. J., Sasaki, S. & Kumazawa, M. *Science* **224**, 749–751 (1984).
4. Duffy, T. S., Zha, C. S., Downs, R. T., Mao, H. K. & Hemley, R. J. *Nature* **378**, 170–173 (1995).
5. Sinogeikin, S. V. & Bass, J. D. *Phys. Earth Planet. Inter.* **120**, 43–62 (2000).
6. Birch, F. *J. Geophys. Res.* **57**, 227–286 (1952).
7. Ringwood, A. E. & Major, A. *Phys. Earth Planet. Inter.* **3**, 89–108 (1970).
8. Liu, L. *Earth Planet. Sci. Lett.* **31**, 200–208 (1976).
9. Li, B. & Zhang, J. *Phys. Earth Planet. Inter.* **151**, 143–154 (2005).
10. Khan, A., Connolly, J. A. D. & Taylor, S. R. *J. Geophys. Res.* **113**, B09308 (2008).
11. Matas, J., Bass, J., Ricard, Y., Mattern, E. & Bukowinski, M. S. T. *Geophys. J. Int.* **170**, 764–780 (2007).
12. Cobden, L. *et al.* *J. Geophys. Res.* **114**, B11309 (2009).
13. Stixrude, L. & Lithgow-Bertelloni, C. *Geophys. J. Int.* **162**, 610–632 (2005).
14. Kennett, B. L. N. & Jackson, I. *Phys. Earth Planet. Inter.* **176**, 98–108 (2009).
15. Murakami, M., Ohishi, Y., Hirao, N. & Hirose, K. *Earth Planet. Sci. Lett.* **277**, 123–129 (2009).
16. Li, B., Woody, K. & Kung, J. J. *Geophys. Res.* **111**, B11206 (2006).

## PLANETARY SCIENCE

## Mercury's mysteries start to unfold

**The origin of the planet Mercury has been a continuing puzzle. Data from NASA's MESSENGER space probe, combined with ground-based observations, are delivering information on the planet's structure and evolution.**

DAVID J. STEVENSON

Writing in *Science*, Smith *et al.*<sup>1</sup> and Zuber *et al.*<sup>2</sup> report analyses of data obtained by NASA's MESSENGER space mission that advance our understanding of Mercury, the innermost of the eight planets in the Solar System. Smith and colleagues determine Mercury's gravity and internal structure, whereas Zuber and co-workers map the topography of the planet's northern hemisphere. Taken together with earlier results on Mercury's magnetic field<sup>3</sup> and variable spin<sup>4</sup>, the new studies suggest that the planet's

structure and evolution are unlike those of any other terrestrial planet or moon in the Solar System.

Because Mercury lies deep within the Sun's gravity potential well, it is hard to get to, and even harder to orbit. MESSENGER is only the second spacecraft to make the journey (after Mariner 10, which arrived in 1974) and the first to enter orbit. It has long been known that Mercury is unusually dense for its size, which means that it must have a large component of iron. Ground-based radar data<sup>4</sup> confirmed the presence of a liquid-iron-rich outer core, suggesting a fully differentiated structure,

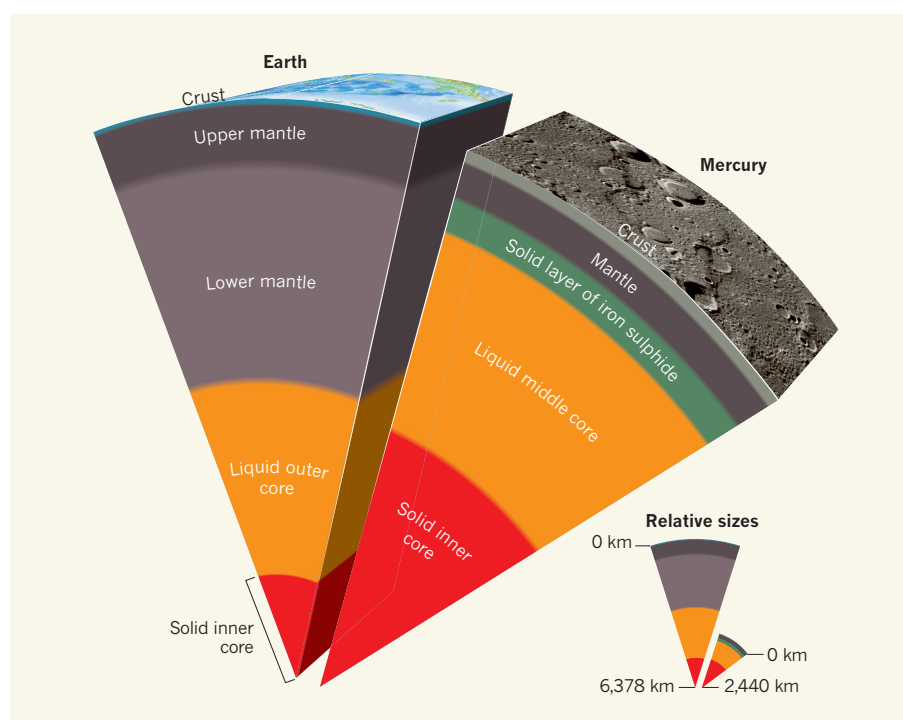
qualitatively like that of Earth. Its magnetic field, known since Mariner 10 but poorly characterized until the advent of data from MESSENGER<sup>3</sup>, suggests an active dynamo, also qualitatively like Earth's.

However, the similarities quickly run out: Mercury's core makes up a far larger fraction of the planet than does Earth's core, and yet the surface magnetic field is much smaller. It has been shown<sup>5</sup> that Mercury's rotational properties, combined with gravity data obtained from precise spacecraft tracking, allow the planet's total moment of inertia to be determined, as well as the moment of inertia of just the outermost solid part, or outer shell. The moment of inertia measures an object's ability to resist changes in its rotation and depends on the object's internal structure and density. The moment-of-inertia values obtained<sup>5</sup> indicate a thin mantle, but also suggest an unusually dense one.

Smith and colleagues' analysis of the MESSENGER data for Mercury's gravity field indicates that the planet's unusually dense outer shell might include a layer of iron sulphide at the base of the mantle (Fig. 1). This solid layer would be part of the planet's outer shell from the point of view of rotational properties, because the material immediately below it is liquid and thus partially decoupled from the libration (pendulum-like motion) that the outer shell undergoes owing to the torque exerted by the Sun on Mercury's distorted shape. It should nonetheless be counted as part of the core because it forms from the core as upward-floating 'snow' as Mercury's interior cools with the ageing Solar System. This unusual structure would be unlike that of any other known planet, but it is permitted by the thermodynamics of iron–sulphur alloys, and might help to explain the planet's small magnetic field.

But is this structure required to explain the data? Perhaps not. The authors acknowledge that a more conventional, Earth-like structure is also conceivable, because MESSENGER's data have an uncertainty in gravity field and planetary obliquity (the tilt of a planet's spin axis relative to an axis perpendicular to its orbit) that implies a rather large margin of error in the value for the total moment of inertia. Even with a more conventional structure, it is interesting that the results suggest that Mercury may have a substantial amount of sulphur relative to iron — perhaps more than Earth. This is surprising, because sulphur is normally considered a volatile element, and yet Mercury is the innermost planet and may therefore have formed at high temperatures, so that the sulphur would be expected never to have been incorporated or to have subsequently been lost.

Meanwhile, Zuber and colleagues' analysis of MESSENGER's altimetry data provides information on Mercury's topography and geological evolution. Flow in the mantle or crust seems likely to have occurred, because



**Figure 1 | Internal structure of Earth and Mercury.** Earth's structure consists of a thin crust overlying a mantle that makes up about two-thirds of the total planetary mass. The iron-rich core has a central, solidified part that may also exist on all terrestrial planets, including Mercury. Smith and colleagues' study<sup>1</sup> suggests that Mercury has a very thin mantle overlying a solid layer of iron sulphide. This layer might have arisen by upward flotation of iron sulphide crystals as the core has cooled, but as the layer contributes to the planet's rotational properties, it counts as part of its outer shell. (Adapted from a figure by NASA/Johns Hopkins Univ. Appl. Phys. Lab./Carnegie Inst. Washington.)

the topography, gravity and images of the planet suggest deformations in its surface over distances of many thousands of kilometres. Although these deformations may be ancient, they post-date much of the planet's impact history and require an explanation based on the planet's internal dynamics. Part of the explanation may lie in the cooling and resulting contraction of Mercury over time.

Zuber *et al.* also show that Mercury has some topographic highs that correlate with gravity highs, but the former are not mere thickenings of crust and are not yet understood. MESSENGER's data allow good regional gravity determination only in the northern hemisphere, but the combination of gravity and topography has allowed the authors to identify large (many tens of kilometres) regional variations in crustal thickness, suggesting a complex volcanic and cooling history.

What does all this say about the origin of Mercury? That is unclear. Two ideas, both of long standing, are in play. One is that Mercury is an outlier in the pattern of terrestrial-planet formation, carrying a memory of thermodynamic equilibrium under conditions different from those thought to have existed when Earth formed: very high temperatures and conditions possibly especially favourable for chemical reduction. This model is well enough defined as to be at least testable, although

whether it can be made consistent with the ideas currently favoured for planet formation, in which a planet's orbital radius is thought to change over time, is unclear.

The alternative theory involves a giant impact in which silicate material is blasted from a proto-Mercury. This is an attractive model because it seems likely that giant impacts are a natural part of planet formation. But it has been insufficiently explored for researchers to judge whether it can meet the challenges of the structure and composition inferred by Smith and colleagues. The European–Japanese mission to Mercury, called BepiColombo, which will launch in 2015 and arrive at the planet in 2022, may help to resolve some of these issues, particularly through its high-quality gravity measurements. ■

**David J. Stevenson** is in the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA.  
e-mail: djs@gps.caltech.edu

1. Smith, D. E. *et al.* *Science* **336**, 214–217 (2012).
2. Zuber, M. T. *et al.* *Science* **336**, 217–220 (2012).
3. Anderson, B. J. *et al.* *Science* **333**, 1859–1862 (2011).
4. Margot, J. L., Peale, S. J., Jurgens, R. F., Slade, M. A. & Holin, I. V. *Science* **316**, 710–714 (2007).
5. Peale, S. J., Phillips, R. J., Solomon, S. C., Smith, D. E. & Zuber, M. T. *Meteorit. Planet. Sci.* **37**, 1269–1283 (2002).



# The translational landscape of mTOR signalling steers cancer initiation and metastasis

Andrew C. Hsieh<sup>1,2\*</sup>, Yi Liu<sup>3\*</sup>, Merritt P. Edlind<sup>1</sup>, Nicholas T. Ingolia<sup>4</sup>, Matthew R. Janes<sup>3</sup>, Annie Sher<sup>1</sup>, Evan Y. Shi<sup>1</sup>, Craig R. Stumpf<sup>1</sup>, Carly Christensen<sup>1</sup>, Michael J. Bonham<sup>5</sup>, Shunyou Wang<sup>3</sup>, Pingda Ren<sup>3</sup>, Michael Martin<sup>3</sup>, Katti Jessen<sup>3</sup>, Morris E. Feldman<sup>6</sup>, Jonathan S. Weissman<sup>6</sup>, Kevan M. Shokat<sup>6</sup>, Christian Rommel<sup>3</sup> & Davide Ruggero<sup>1</sup>

The mammalian target of rapamycin (mTOR) kinase is a master regulator of protein synthesis that couples nutrient sensing to cell growth and cancer. However, the downstream translationally regulated nodes of gene expression that may direct cancer development are poorly characterized. Using ribosome profiling, we uncover specialized translation of the prostate cancer genome by oncogenic mTOR signalling, revealing a remarkably specific repertoire of genes involved in cell proliferation, metabolism and invasion. We extend these findings by functionally characterizing a class of translationally controlled pro-invasion messenger RNAs that we show direct prostate cancer invasion and metastasis downstream of oncogenic mTOR signalling. Furthermore, we develop a clinically relevant ATP site inhibitor of mTOR, INK128, which reprograms this gene expression signature with therapeutic benefit for prostate cancer metastasis, for which there is presently no cure. Together, these findings extend our understanding of how the 'cancerous' translation machinery steers specific cancer cell behaviours, including metastasis, and may be therapeutically targeted.

It is unknown whether specialized networks of translationally controlled mRNAs can direct cancer initiation and progression, thereby mirroring cooperativity that has mainly been observed at the level of transcriptional control. This is an important question, as key oncogenic signalling molecules, such as the mTOR kinase, directly regulate the activity of general translation factors<sup>1,2</sup>. Downstream of the phosphatidylinositol-3-OH kinase (PI(3)K)–AKT signalling pathway, mTOR assembles with either raptor or rictor to form two distinct complexes: mTORC1 and mTORC2 (refs 3, 4). The major regulators of protein synthesis downstream of mTORC1 are 4EBP1 (also called EIF4EBP1) and p70S6K1/2 (refs 1, 2). 4EBP1 negatively regulates eIF4E, a key rate-limiting initiation factor for cap-dependent translation. Phosphorylation of 4EBP1 by mTORC1 leads to its dissociation from eIF4E, allowing translation initiation complex formation at the 5' end of mRNAs<sup>5</sup>. The mTOR-dependent phosphorylation of p70S6K1/2 also promotes translation initiation as well as elongation<sup>6</sup>. At a genome-wide level, it remains poorly understood whether and how activation of these regulators of protein synthesis may produce specific changes in gene expression networks that direct cancer development. Here we use a powerful new technology known as ribosome profiling to delineate the translational landscape of the cancer genome at a codon-by-codon resolution upon pharmacological inhibition of mTOR<sup>7</sup>. Our findings provide genome-wide characterization of translationally controlled mRNAs downstream of oncogenic mTOR signalling and delineate their functional roles in cancer development. Moreover, we determine the efficacy of a novel clinically relevant mTOR inhibitor that we developed, INK128, which specifically targets this cancer program.

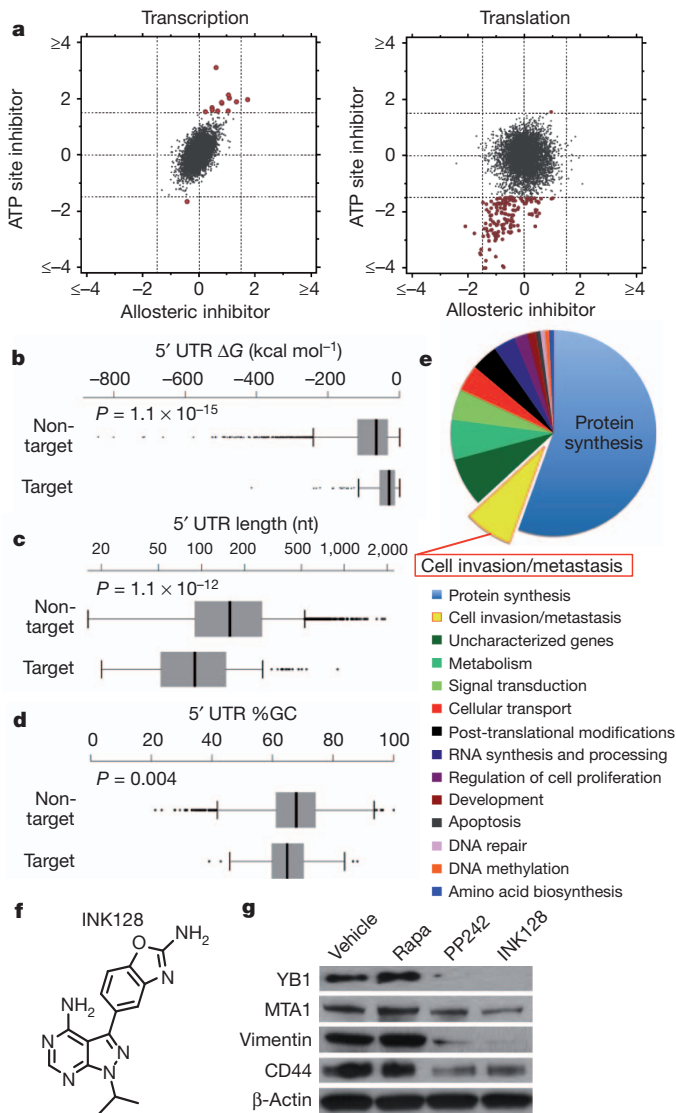
## Ribosome profiling of the prostate cancer genome

mTOR is deregulated in nearly 100% of advanced human prostate cancers<sup>8</sup>, and genetic findings in mouse models implicate mTOR hyperactivation in prostate cancer initiation<sup>9–11</sup>. Given the critical role for mTOR in prostate cancer, we used PC3 human prostate cancer cells, where mTOR is constitutively hyperactivated, to delineate translationally controlled gene expression networks upon complete or partial mTOR inhibition. We optimized ribosome profiling to assess quantitatively ribosome occupancy genome-wide in cancer cells<sup>7</sup>. In brief, ribosome-protected mRNA fragments were deep-sequenced to determine the number of ribosomes engaged in translating specific mRNAs (Supplementary Fig. 1a and Methods). Treatment of PC3 cells with PP242 (refs 12, 13), an mTOR ATP site inhibitor, significantly inhibits the activity of the three primary downstream mTOR effectors 4EBP1, p70S6K1/2 and AKT. On the contrary, rapamycin, an allosteric mTOR inhibitor, only blocks p70S6K1/2 activity in these cells (Supplementary Fig. 1b). We used short 3-h drug treatments, which precede alterations in *de novo* protein synthesis, to capture direct changes in mTOR-dependent gene expression by ribosome profiling and to minimize compensatory feedback mechanisms (Supplementary Fig. 1c–f).

Ribosome profiling revealed 144 target mRNAs selectively decreased at the translational level upon PP242 treatment ( $\log_2 \leq -1.5$  (false discovery rate  $< 0.05$ )) compared to rapamycin treatment, with limited changes in transcription (Fig. 1a and Supplementary Figs 2a, b and 3–10). The fact that at this time point rapamycin treatment did not markedly affect gene expression is consistent with incomplete allosteric inhibition of mTOR activity (Supplementary Fig. 1b). By

<sup>1</sup>School of Medicine and Department of Urology, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California 94158, USA. <sup>2</sup>Division of Hematology/Oncology and Department of Internal Medicine, University of California, San Francisco, California 94143, USA. <sup>3</sup>Intellikine Inc., La Jolla, California 92037, USA. <sup>4</sup>Carnegie Institution for Science, Baltimore, Maryland 21218, USA. <sup>5</sup>Department of Pathology, University of California, San Francisco, California 94143, USA. <sup>6</sup>Howard Hughes Medical Institute, Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, USA.

\*These authors contributed equally to this work.



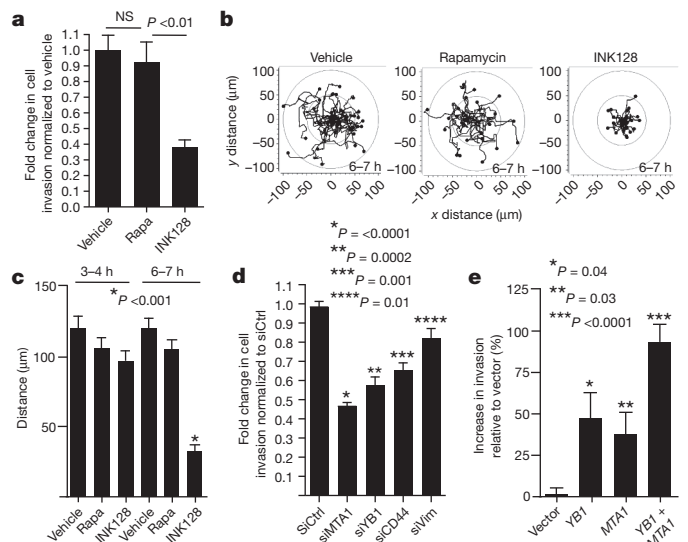
**Figure 1 | Ribosome profiling reveals mTOR-dependent specialized translational control of the prostate cancer genome.** **a**, Representative comparison of mRNA abundance and translational efficiency after a 3-h treatment with an ATP site inhibitor (PP242) versus an allosteric inhibitor (rapamycin). **b–d**, Free energy, length and percentage G + C content of the 5' UTRs of mTOR target versus non-target mRNAs (error bars indicate range, non-target  $n = 5,022$ , target  $n = 144$ , two-sided Wilcoxon). **e**, Functional classification of translationally regulated mTOR-responsive mRNAs. **f**, Chemical structure of INK128. **g**, Representative western blot from three independent experiments of mTOR-sensitive invasion genes in PC3 cells after a 48-h drug treatment. Rapa, rapamycin.

monitoring footprints of translating 80S ribosomes, our findings show that the effects of PP242 are largely at the level of translation initiation and not elongation (Supplementary Fig. 3). It has been proposed that mRNAs translationally regulated by mTOR may contain long 5' untranslated regions (5' UTRs) with complex RNA secondary structures. On the contrary, ribosome profiling revealed that mTOR-responsive 5' UTRs possess less complex features (Fig. 1b–d), providing a unique data set to investigate the nature of regulatory elements that render these mRNAs mTOR-sensitive. It has been previously shown that some mTOR translationally regulated mRNAs, most notably those involved in protein synthesis, possess a 5' terminal oligopyrimidine tract (5' TOP)<sup>14,15</sup> that is regulated by distinct *trans*-acting factors<sup>16,17</sup>. Of the 144 mTOR-sensitive target genes, 68% possess a 5' TOP. However, as the 5' TOP is not present in all mTOR-sensitive mRNAs, we next asked whether other 5' UTR

consensus sequences may exist. Strikingly, 63% of mTOR target mRNAs possess what we have termed a pyrimidine-rich translational element (PRTE) within their 5' UTRs ( $P = 3.2 \times 10^{-11}$ ). This element, unlike the 5' TOP sequence, consists of an invariant uridine at position 6 flanked by pyrimidines and, importantly, does not reside at position +1 of the 5' UTR (Supplementary Figs 2c and 7). We found that 89% of the mTOR-responsive genes possess a PRTE and/or 5' TOP, making the presence of one or both sequences a strong predictor for mTOR sensitivity (Supplementary Figs 2d and 7). Notably, mRNA isoforms arising from distinct transcription start sites may possess both a 5' TOP and a PRTE. Moreover, given the significant number of mRNAs that contain both the PRTE and 5' TOP, a functional interplay may exist between these regulatory elements. Future studies are required to determine the regulatory logic for how these sequences either independently or coordinately confer mTOR responsiveness. Multiple *cis*-acting elements within specific 5' UTRs could reflect regulation by distinct mTOR effectors. For example, our findings show that the PRTE imparts translational control specificity to 4EBP1 activity (see below).

Surprisingly, mTOR-sensitive genes stratify into unique functional categories that may promote cancer development and progression, including cellular invasion ( $P = 0.009$ ), cell proliferation ( $P = 0.04$ ), metabolism ( $P = 0.0002$ ) and regulators of protein modification ( $P = 0.01$ ) (Fig. 1e). The largest fraction of mTOR-responsive mRNAs cluster into a node consisting of key components of the translational apparatus: 70 ribosomal proteins, 6 elongation factors, and 4 translation initiation factors ( $P = 7.5 \times 10^{-82}$ ) (Fig. 1e and Supplementary Fig. 5). Therefore, this class of mTOR-responsive mRNAs may represent an important regulon that sustains the elevated protein synthetic capacity of cancer cells.

Notably, the second largest node of mTOR translationally regulated genes comprises bona fide cell invasion and metastasis mRNAs and putative regulators of this process (Fig. 1e). This group includes YB1 (Y-box binding protein 1; also called YBX1), vimentin, MTA1 (metastasis associated 1) and CD44 (Supplementary Fig. 11a). YB1



**Figure 2 | mTOR promotes prostate cancer cell migration and invasion through a translationally regulated gene signature.** **a**, Matrigel invasion assay in PC3 cells: 6-h pre-treatment followed by 6 h of cell invasion ( $n = 6$ , ANOVA). **b**, **c**, Migration patterns and average distance travelled by GFP-labelled PC3 cells during hours 3–4 and 6–7 of drug treatment ( $n = 34$  cells per condition, ANOVA). **d**, Matrigel invasion assay in PC3 cells after 48 h of knockdown of YB1, MTA1, CD44, or vimentin followed by 24 h of cell invasion ( $n = 7$ , *t*-test). **e**, Matrigel invasion assay in BPH-1 cells after 48 h of overexpression of YB1 and/or MTA1, followed by cell invasion for 24 h ( $n = 7$ , *t*-test). Rapa, rapamycin. All data represent mean  $\pm$  s.e.m. NS, not statistically significant.



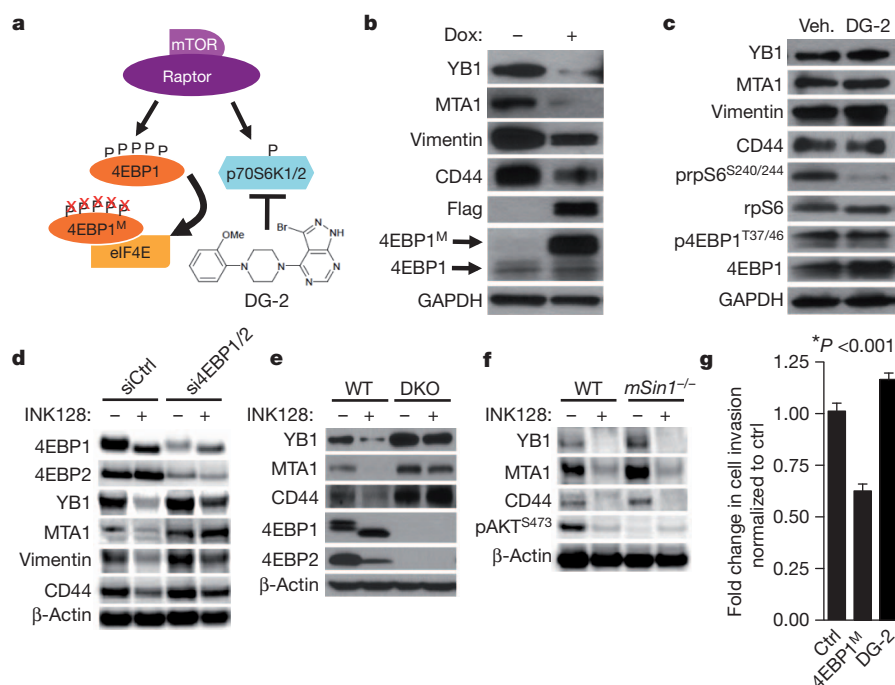
regulates the post-transcriptional expression of a network of invasion genes<sup>18</sup>. Vimentin, an intermediate filament protein, is highly upregulated during the epithelial-to-mesenchymal transition associated with cellular invasion<sup>19</sup>. MTA1, a putative chromatin-remodelling protein, is overexpressed in invasive human prostate cancer<sup>20</sup> and has been shown to drive cancer metastasis by promoting neoangiogenesis<sup>21</sup>. CD44 is commonly overexpressed in tumour-initiating cells and is implicated in prostate cancer metastasis<sup>22</sup>. Consistent with their status as mTOR sensitive genes, *YB1*, vimentin, *MTA1* and *CD44* all possess a PRTE (Supplementary Fig. 5). Vimentin and *CD44* also possess a 5' TOP (Supplementary Fig. 7). To test the functional role of the PRTE in mediating translational control, we mutated the PRTE within the 5' UTR of *YB1*, which rendered the *YB1* 5' UTR insensitive to inhibition by 4EBP1 (Supplementary Fig. 11b). These findings highlight a novel *cis*-regulatory element that may modulate translational control of subsets of mRNAs upon mTOR activation. Moreover, ribosome profiling reveals unexpected transcript-specific translational control, mediated by oncogenic mTOR signalling, including a distinct set of pro-invasion and metastasis genes.

### Translation of pro-invasion mRNAs by mTOR

We next extended the use of the mTOR pharmacological tools used in ribosome profiling towards functional characterization of the newly identified mTOR-sensitive cell invasion gene signature. To this end, we developed a new clinical-grade mTOR ATP site inhibitor, INK128, derived from the PP242 chemical scaffold (Fig. 1f). In brief, a structure-guided optimization of pyrazolopyrimidine derivatives was performed (see INK128 chemical synthesis in Supplementary Information) that improved oral bioavailability while retaining mTOR kinase potency and selectivity. INK128 was selected for clinical studies on the basis of its high potency (1.4 nM inhibition constant ( $K_i$ )), selectivity for mTOR, low molecular mass, and favourable pharmaceutical properties (Supplementary Figs 12 and 13).

Using either PP242 or INK128, we observed a selective decrease in the expression of *YB1*, *MTA1*, vimentin and *CD44* at the protein but not transcript level in PC3 cells starting at 6 h of treatment, which precedes any decrease in *de novo* protein synthesis (Fig. 1g and Supplementary Figs 1c, d, 14 and 15). In contrast, rapamycin treatment did not alter their expression (Fig. 1g and Supplementary Fig. 14a). Similar findings were observed using a broad panel of metastatic cell lines of distinct histological origins (Supplementary Fig. 16). The four-gene invasion signature is positively regulated by mTOR hyperactivation, as silencing PTEN expression increased their protein but not mRNA expression levels (Supplementary Fig. 17). We next investigated the effects of mTOR ATP site inhibitors on prostate cancer cell migration and invasion. We found that INK128, but not rapamycin, decreases the invasive potential of PC3 prostate cancer cells (Fig. 2a). Furthermore, INK128 inhibits cancer cell migration starting at 6 h of treatment, precisely correlating with when decreases in the expression of pro-invasion genes are evident, but preceding any changes in the cell cycle or overall global protein synthesis (Fig. 2b, c, and Supplementary Figs 1c, e, f, 14b and 18).

Among the genes comprising the pro-invasion signature, *YB1* has been shown to act directly as a translation factor that controls expression of a larger set of genes involved in breast cancer cell invasion<sup>18</sup>. Notably, *YB1* translationally regulated target mRNAs, including *SNAIL1* (also called *SNAIL*), *LEF1* and  *Twist1*, decreased at the protein but not transcript level upon *YB1* knockdown in PC3 cells (Supplementary Figs 19 and 20). To determine the functional role of *YB1* in prostate cancer cell invasion, we silenced *YB1* gene expression in PC3 cells, and observed a 50% reduction in cell invasion (Fig. 2d). Similarly, knockdown of *MTA1*, *CD44*, or vimentin also inhibited prostate cancer cell invasion (Fig. 2d and Supplementary Fig. 19). These mTOR target mRNAs may be sufficient to endow primary prostate cells with invasive features, as overexpression of *YB1* and/or *MTA1* (Supplementary Fig. 21a) in BPH-1 cells, an untransformed



**Figure 3 | The 4EBP1-eIF4E axis controls the post-transcriptional expression of mTOR-sensitive invasion genes.** **a**, Schematic of the pharmacogenetic strategy to inhibit p70S6K1/2 or eIF4E hyperactivation. **b**, Representative western blot from three independent experiments of PC3 4EBP1<sup>M</sup> cells after 48-h doxycycline induction of 4EBP1<sup>M</sup>. **c**, Representative western blot from three independent experiments of PC3 cells after 48-h DG-2 treatment. **d**, Representative western blot from three independent experiments of PC3 cells after 48 h of 4EBP1/4EBP2 knockdown followed by 24-h INK128

treatment (see quantification of independent experiments in Supplementary Fig. 23a). **e**, Representative western blot from three independent experiments of wild type (WT) and 4EBP1/4EBP2 double knockout (DKO) MEFs treated with INK128 for 24 h. **f**, Representative western blot from two independent experiments of wild-type and *mSin1*<sup>-/-</sup> (also called *Mapkap1*<sup>tm1Bisu</sup>) MEFs after 24-h INK128 treatment. **g**, Matrigel invasion assay upon 48-h doxycycline induction of 4EBP1<sup>M</sup>, or treatment with DG-2 compared to control ( $n = 6$  per condition,  $t$ -test). All data represent mean  $\pm$  s.e.m.

prostate epithelial cell line, increased the invasive capacity of these cells in an additive manner (Fig. 2e). Notably, the effects of YB1 and MTA1 on cell invasion are independent from any effect on cell proliferation in both knockdown or overexpression studies (Supplementary Fig. 21b, c). Therefore, translational control of pro-invasion mRNAs by oncogenic mTOR signalling alters the ability of epithelial cells to migrate and invade, a key feature of cancer metastasis.

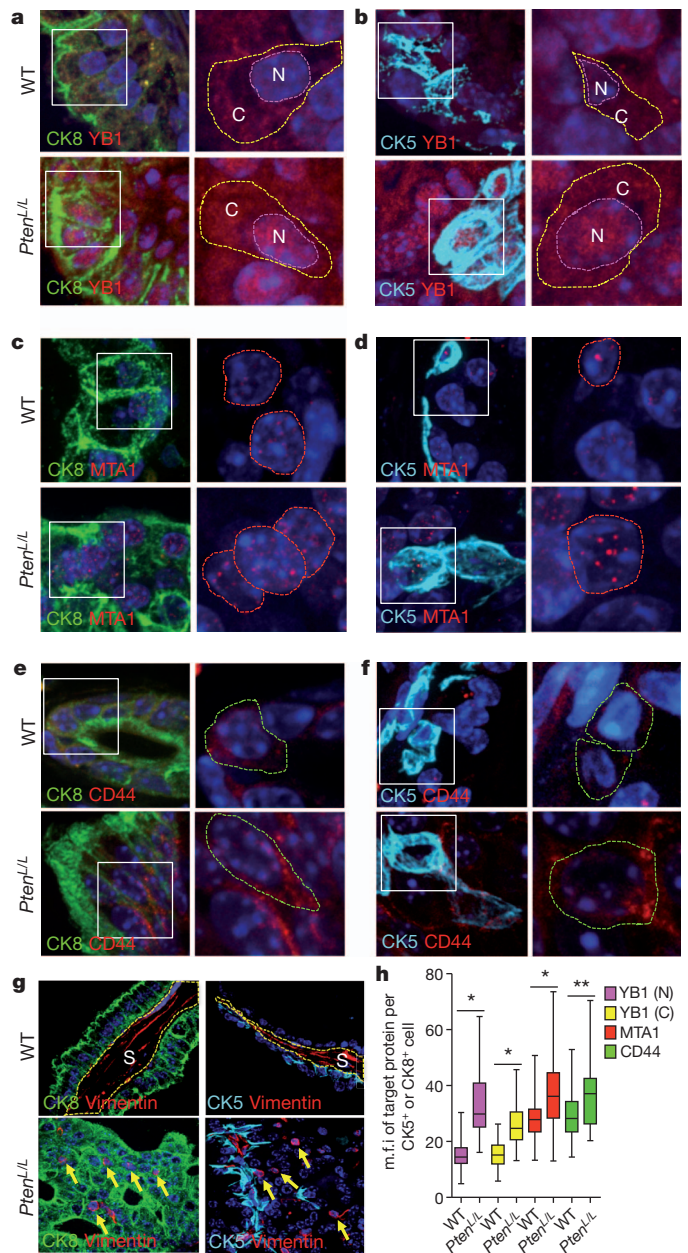
### Dissecting mTOR translational effectors

We sought to determine the molecular mechanism by which pro-invasion genes are regulated at the translational level and why these mRNAs are sensitive to INK128 but not rapamycin. To this end, we investigated whether the translational regulators downstream mTORC1, 4EBP1 and/or p70S6K1/2, control the expression of these mTOR-sensitive targets. We generated a human prostate cancer cell line that stably expresses a doxycycline-inducible dominant-negative mutant of 4EBP1 (4EBP1<sup>M</sup>) (Fig. 3a)<sup>13</sup>. This mutant binds to eIF4E, decreasing its hyperactivation without inhibiting general mTORC1 function (Supplementary Fig. 22a). Notably, expression of 4EBP1<sup>M</sup> does not alter global protein synthesis (Supplementary Fig. 22b), probably because endogenous 4EBP1 and 4EBP2 proteins retain their ability to bind to eIF4E (Supplementary Fig. 22c)<sup>13</sup>. Upon induction of 4EBP1<sup>M</sup>, YB1, vimentin, CD44 and MTA1 decrease at the protein but not mRNA level, whereas pharmacological inhibition of p70S6K1/2 with DG-2 (ref. 23) had no effect (Fig. 3b, c and Supplementary Fig. 22d). Next, we tested whether INK128 decreases expression of the four invasion genes through the 4EBP–eIF4E axis. Notably, knockdown of 4EBP1 and 4EBP2 in PC3 cells or using 4EBP1 and 4EBP2 double knockout mouse embryonic fibroblasts (MEFs)<sup>24</sup> reduced the ability of INK128 to decrease expression of these pro-invasion mRNAs (Fig. 3d, e and Supplementary Fig. 23). Furthermore, ablation of mTORC2 activity<sup>25</sup> had no effect on the expression of these mRNAs or responsiveness to INK128 (Fig. 3f and Supplementary Fig. 24a–c). Next, we determined the effect of 4EBP1<sup>M</sup> on human prostate cancer cell invasion. The expression of 4EBP1<sup>M</sup> resulted in a significant decrease in prostate cancer cell invasion without affecting the cell cycle, whereas DG-2 had no effect (Fig. 3g and Supplementary Fig. 24d). These findings demonstrate that eIF4E hyperactivation downstream of oncogenic mTOR regulates translational control of the pro-invasion mRNAs and provides an explanation for the selective targeting of this gene signature by mTOR ATP site inhibitors.

### Examining cell invasion networks in vivo

Both CK5<sup>+</sup> and CK8<sup>+</sup> prostate epithelial cells have been implicated in the initiation of prostate cancer upon loss of PTEN<sup>26,27</sup>. *Pten*<sup>loxP/loxP</sup>; *Pb-cr* (*Pten*<sup>L/L</sup>) mice are an ideal model of prostate cancer because they display distinct stages of cancer development (prostatic intraepithelial neoplasia, invasive adenocarcinoma, and metastasis)<sup>28</sup>. However, the expression patterns of YB1, vimentin, CD44 and MTA1 in prostate basal (CK5<sup>+</sup>) and luminal (CK8<sup>+</sup>) epithelial cells have not been characterized. We therefore analysed their expression patterns in the *Pten*<sup>L/L</sup> prostate cancer mouse model, where mTOR is constitutively hyperactivated<sup>9,28</sup>. We found that YB1 localizes to the cytoplasm and nucleus of CK5<sup>+</sup> and CK8<sup>+</sup> prostate epithelial cells, consistent with its ability to shuttle between the two cellular compartments (Fig. 4a, b and Supplementary Fig. 25a, b)<sup>18,29</sup>. MTA1 expression is exclusively nuclear in both cell types (Fig. 4c, d). Of note, CD44, together with other cell-surface markers, has been used to isolate a rare prostate stem-cell population<sup>30</sup>. We observed expression of CD44 within a subset of CK5<sup>+</sup> and CK8<sup>+</sup> epithelial cells (Fig. 4e, f). In contrast, vimentin is not detected in either cell type (Fig. 4g). We next determined the impact of mTOR hyperactivation on the expression pattern of the pro-invasion gene signature. YB1, MTA1 and CD44 protein, but not transcript, levels were significantly increased in both *Pten*<sup>L/L</sup> luminal and basal epithelial cells compared to wild type (Fig. 4h and Supplementary Fig. 25c–e). Interestingly, a subset of *Pten*<sup>L/L</sup> luminal epithelial cells ectopically

expresses vimentin at aberrantly high levels, with a perinuclear distribution (Fig. 4g and Supplementary Fig. 25f, g) suggesting that these cells may have acquired some mesenchymal-like features. Consistent with these findings, perinuclear vimentin localization is associated with invasive features of human prostate cancer cells<sup>31</sup> and changes



**Figure 4 | mTOR hyperactivation augments translation of YB1, MTA1, CD44 and vimentin mRNAs in a subset of pre-invasive prostate cancer cells in vivo.** Left: immunofluorescent images of CK8/DAPI or CK5/DAPI with YB1 (a, b), MTA1 (c, d), or CD44 (e, f) co-staining in 14-month-old wild-type and *Pten*<sup>L/L</sup> mouse prostate epithelial cells. White boxes outline the area magnified in the right panel. Right: magnified immunofluorescent images of YB1 (a, b), MTA1 (c, d) and CD44 (e, f) co-stained with DAPI in wild-type and *Pten*<sup>L/L</sup> mouse prostate epithelial cells. Dotted lines encircle the cytoplasm (C) and/or the nucleus (N). g, Representative immunofluorescent images of CK5 or CK8 co-staining with vimentin in 14-month-old wild-type and *Pten*<sup>L/L</sup> mouse prostate epithelial cells. S, stroma; yellow arrows indicate perinuclear vimentin. h, Box plot of YB1 (N = nuclear, C = cytoplasmic), MTA1 and CD44 mean fluorescence intensity (m.f.i.) per CK5<sup>+</sup> or CK8<sup>+</sup> prostate epithelial cell in wild-type and *Pten*<sup>L/L</sup> mice (three mice per arm,  $n = 43$ –303 cells quantified per target gene, error bars indicate range (see Supplementary Fig. 25b); \* $P < 0.0001$ , \*\* $P = 0.0004$ ,  $t$ -test).

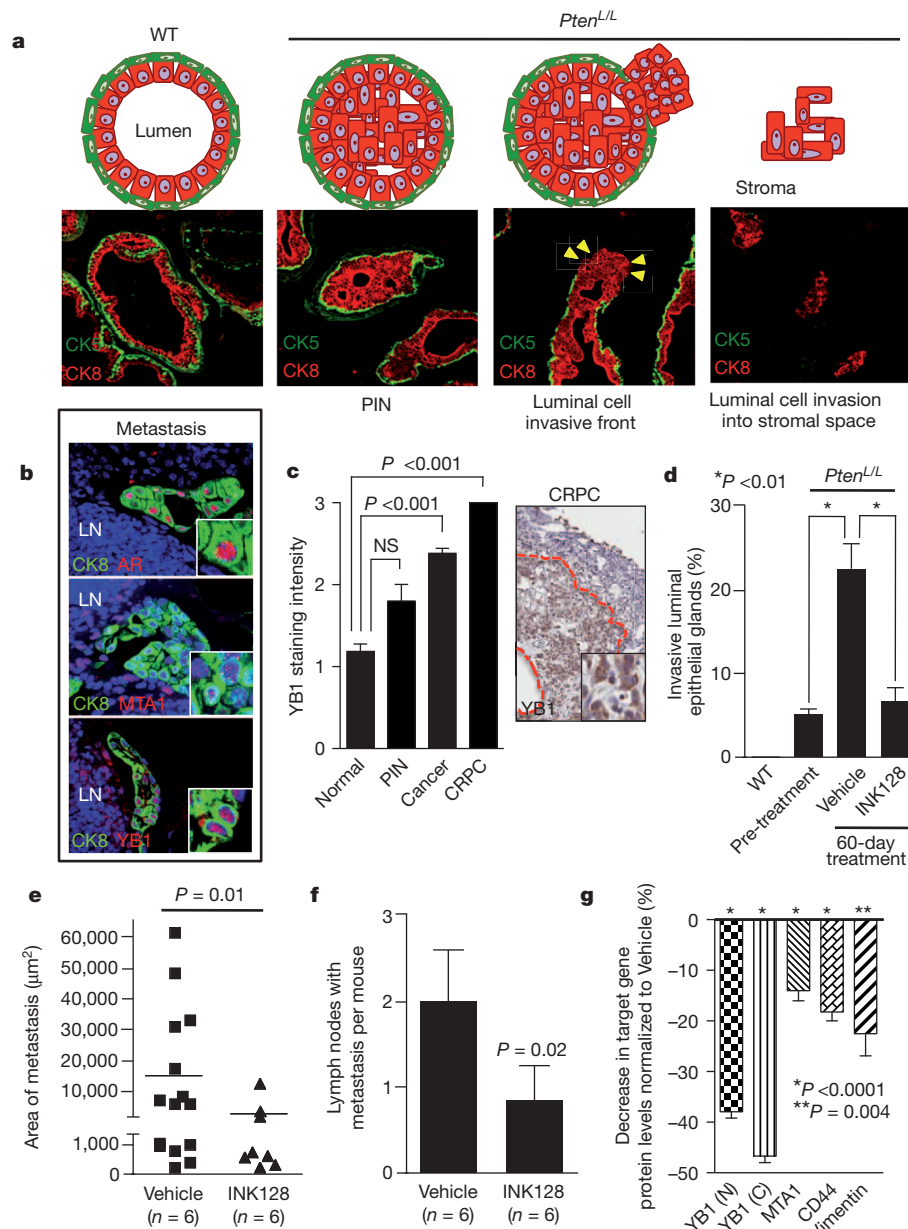


in cell polarity in actively moving fibroblasts<sup>32</sup>. These studies reveal a unique, translationally controlled signature of gene expression downstream of mTOR hyperactivation in a cancer-initiating subset of prostate epithelial cells.

### Targeting prostate cancer metastasis

The most significant pre-clinical extension of this work would be to determine the therapeutic benefit of INK128 in reprogramming expression of the mTOR-dependent pro-invasion gene signature and prostate cancer metastasis directly *in vivo*. This is underscored by the clinical inefficacy of allosteric mTOR inhibition towards the

lethal form of metastatic human prostate cancer<sup>33,34</sup>. Importantly, in our preclinical trial of RAD001 (rapalog) versus INK128 in *Pten*<sup>L/L</sup> mice, 4EBP1 and p70S6K1/2 phosphorylation was completely restored to wild-type levels after treatment with INK128, whereas RAD001 only decreased p70S6K1/2 phosphorylation levels (Supplementary Fig. 26a, b). We next determined the cellular consequences of complete versus partial mTOR inhibition during distinct stages of prostate cancer. INK128 treatment resulted in a 50% decrease in prostatic intraepithelial neoplasia (PIN) lesions in *Pten*<sup>L/L</sup> mice that was associated with decreased proliferation and a tenfold increase in apoptosis (Supplementary Fig. 26d–f). Notably,



**Figure 5 | Complete mTOR inhibition by INK128 treatment prevents prostate cancer invasion and metastasis *in vivo*.** **a**, Diagram and images of normal prostate gland, pre-invasive PIN and invasive prostate cancer. CK8/CK5, luminal/basal epithelial cells, respectively. Yellow arrowheads indicate invasive front. **b**, Immunofluorescent images of 14-month-old *Pten*<sup>L/L</sup> lymph node (LN) metastasis co-stained with CK8/androgen receptor (AR), CK8/YB1 and CK8/MTA1. **c**, Left: human tissue microarray of YB1 protein levels in normal ( $n = 59$ ), PIN ( $n = 5$ ), cancer ( $n = 99$ ) and CRPC ( $n = 3$ ) (ANOVA). Right: immunohistochemistry of YB1 in human CRPC demarcated by the red line (inset shows nuclear and cytoplasmic YB1). **d**, Quantification of invasive

prostate glands in wild-type and *Pten*<sup>L/L</sup> mice before (12-months old) and after (14-months old) 60 days of INK128 treatment ( $n = 6$  mice per arm, ANOVA). **e**, **f**, Area and number of CK8/AR<sup>+</sup> metastases in draining lymph nodes in 14-month-old *Pten*<sup>L/L</sup> mice after 60 days of INK128 treatment ( $n = 6$  mice per arm, *t*-test). **g**, Percentage decrease of YB1 (N = nuclear, C = cytoplasmic), MTA1, CD44, or vimentin protein levels (determined by quantitative immunofluorescence, Supplementary Fig. 25b) in CK8<sup>+</sup> or CK5<sup>+</sup> prostate cells (CK8<sup>+</sup> only for vimentin) in INK128-treated 14-month-old *Pten*<sup>L/L</sup> mice normalized to vehicle-treated mice ( $n = 3$  mice per arm, *t*-test). All data represent mean  $\pm$  s.e.m.

the unique cytotoxic properties of INK128 treatment in *Pten*<sup>L/L</sup> mice were evidenced by a marked reduction in prostate cancer volume. In addition, and consistent with these findings, INK128 induced programmed cell death in multiple cancer cell lines (Supplementary Fig. 27a, b). In contrast, RAD001 treatment mainly had cytostatic effects leading to only partial regression of PIN lesions associated with a limited decrease in cell proliferation and no significant effect on apoptosis (Supplementary Fig. 26c–f).

We extended the preclinical trial by examining the effects of INK128 treatment on the pro-invasion gene signature and prostate cancer metastasis, which is incurable and the primary cause of patient mortality. Cell invasion is the critical first step in metastasis, required for systemic dissemination. In *Pten*<sup>L/L</sup> mice after the onset of PIN, a subset of prostate glands show characteristics of luminal epithelial cell invasion by 12 months (Fig. 5a and Supplementary Fig. 27c)<sup>28</sup>. After 12 months of age, *Pten*<sup>L/L</sup> mice develop lymph-node metastases and these cells maintain strong YB1 and MTA1 expression (Fig. 5b). We further extended these findings directly to human prostate cancer patient specimens, observing that YB1 expression levels increase in a stepwise fashion from normal prostate to castration-resistant prostate cancer (CRPC), an advanced form of the disease associated with increased metastatic potential (Fig. 5c). MTA1 levels exhibit similar increases<sup>20</sup>. In human prostate cancer, high-grade primary tumours that display invasive features are more likely to develop systemic metastasis than low-grade non-invasive tumours<sup>35,36</sup>. Remarkably, treatment with INK128 completely blocked the progression of invasive prostate cancer locally in the prostate gland, and profoundly inhibited the total number and size of distant metastases (Fig. 5d–f). This was associated with a marked decrease in the expression of YB1, vimentin, CD44 and MTA1 at the protein, but not transcript, level in specific epithelial cell types within pre-invasive PIN lesions in *Pten*<sup>L/L</sup> mice (Fig. 5g and Supplementary Fig. 25c). Together, these findings reveal an unexpected role for oncogenic mTOR signalling in control of a pro-invasion translational program that, along with the lethal metastatic form of prostate cancer, can be efficiently targeted with clinically relevant mTOR ATP site inhibitors.

## Discussion

Here we used ribosome profiling to generate a comprehensive map of translationally controlled mTOR targets in cancer that surprisingly stratify into specific cellular processes including proliferation, metabolism, protein synthesis and invasion (Fig. 1e). The effects of this translational control program are probably broad, converging on many subclasses of mRNAs that may cooperate towards distinct steps in cancer development and therapeutic response. This is supported by our *in vivo* findings where we developed a novel clinically relevant mTOR inhibitor, INK128, that significantly abrogates multiple aspects of prostate cancer development by inducing apoptosis as well as inhibiting cell proliferation, invasion and metastasis (Fig. 5d–g and Supplementary Fig. 26c–f). The superiority of INK128 as an mTOR inhibitor is also evident in its ability to reprogram the mTOR oncogenic translational program in prostate cancer, which is not achieved by rapalog treatment. Of note, however, the sensitivity of cells from distinct histological origins to ATP site versus allosteric inhibitors of mTOR may differ. For example, the Jurkat lymphoid cell line is particularly sensitive to rapamycin treatment<sup>37</sup>.

One of the most novel nodes of mTOR translationally controlled genes are those that cooperatively control, at least in part, the cellular invasive features of human prostate cancer cells (Figs 1g, 2 and 3b, g). Translational control of these mRNAs relies on the 4EBP1–eIF4E axis and is thereby specifically druggable with potent mTOR ATP site inhibitors, which, unlike rapamycin, target mTOR-dependent 4EBP1 phosphorylation (Figs 1g, 3d, e and 5g, and Supplementary Figs 1b, 23 and 26b). This has significant therapeutic implications not only for advanced prostate cancer but also for multiple metastatic cancers where we show that translational control of pro-invasion

mRNAs is sensitive to ATP site inhibitors of mTOR (Supplementary Fig. 16). Thereby, these studies link translational regulation to the poorly understood mechanisms underlying cancer metastasis. Intriguingly, various components of the translation machinery, including oncogenic eIF4E<sup>38</sup>, localize to the leading edge of migrating fibroblasts<sup>39</sup>. This may allow spatially controlled translation of mRNAs important for cell migration, providing a rapid and specific response in transducing a migration program that could be co-opted at the invasive edge of metastatic cancer cells. Together, these studies reveal that the ability of mTOR to phosphorylate general translation factors results in exquisite transcript-specific translational control of key mRNAs that may cooperate in distinct steps of cancer initiation and progression, with significant implications for therapeutic intervention.

## METHODS SUMMARY

**Mice.** *Pten*<sup>loxP/loxP</sup> and *Pb-cre* mice were obtained from Jackson Laboratories and Mouse Models of Human Cancers Consortium (MMHCC) and maintained in the C57BL/6 background.

**Ribosome profiling.** PC3 lysates were subjected to ribosome footprinting by nuclease treatment. Ribosome-protected and alkaline digested poly(A) mRNA fragments were purified and deep sequencing libraries were generated. Ribosome footprint and RNA-seq sequencing reads were aligned against a library of transcripts from the UCSC Known Genes database GRCh37/hg19. Read density profiles were constructed for the canonical transcript of each gene. The average read density per codon was computed for the coding sequence of each transcript. Average read density was used to determine mRNA abundance (RNA-seq reads), ribosome occupancy of mRNAs (foot print reads), and translational efficiency (foot print reads/RNA-seq reads).

**Immunofluorescence.** Paraffin-embedded mouse prostates and lymph nodes were deparaffinized and rehydrated using CitriSolv (Fisher) and serial ethanol washes. Antigen unmasking was performed using Citrate pH 6 (Vector Labs). Sections were blocked in 5% goat serum, 1% BSA in TBS. Various primary antibodies were used at dilutions between 1:50 and 1:500 (see Methods), followed by incubation with appropriate conjugated secondary antibodies. Samples were mounted with DAPI Hardset Mounting Medium (Vector Lab). A Zeiss Spinning Disc confocal (Zeiss, CSU-X1) was used to image the tissues. Individual cells were quantified for mean fluorescence intensity using the Axiovision (Zeiss, Release 4.8) densitometric tool.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 14 July 2011; accepted 3 February 2012.

Published online 22 February 2012.

1. Brown, E. J. *et al.* Control of p70 s6 kinase by kinase activity of FRAP *in vivo*. *Nature* **377**, 441–446 (1995).
2. Gingras, A. C., Kennedy, S. G., O'Leary, M. A., Sonenberg, N. & Hay, N. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev.* **12**, 502–513 (1998).
3. Kim, D. H. *et al.* mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* **110**, 163–175 (2002).
4. Sarbassov, D. D. *et al.* Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. *Curr. Biol.* **14**, 1296–1302 (2004).
5. Gingras, A. C., Raught, B. & Sonenberg, N. Regulation of translation initiation by FRAP/mTOR. *Genes Dev.* **15**, 807–826 (2001).
6. Ruvinsky, I. & Meyuhas, O. Ribosomal protein S6 phosphorylation: from protein synthesis to cell size. *Trends Biochem. Sci.* **31**, 342–348 (2006).
7. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
8. Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).
9. Nardella, C. *et al.* Differential requirement of mTOR in postmitotic tissues and tumorigenesis. *Sci. Signal.* **2**, ra2 (2009).
10. Guertin, D. A. *et al.* mTOR complex 2 is required for the development of prostate cancer induced by Pten loss in mice. *Cancer Cell* **15**, 148–159 (2009).
11. Furic, L. *et al.* eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc. Natl Acad. Sci. USA* **107**, 14134–14139 (2010).
12. Feldman, M. E. *et al.* Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biol.* **7**, e38 (2009).
13. Hsieh, A. C. *et al.* Genetic dissection of the oncogenic mTOR pathway reveals druggable addiction to translational control via 4EBP-eIF4E. *Cancer Cell* **17**, 249–261 (2010).



14. Tang, H. *et al.* Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Mol. Cell. Biol.* **21**, 8671–8683 (2001).
15. Meyuhas, O. Synthesis of the translational apparatus is regulated at the translational level. *Eur. J. Biochem.* **267**, 6321–6330 (2000).
16. Crosio, C., Boyl, P. P., Loreni, F., Pierandrei-Amaldi, P. & Amaldi, F. La protein has a positive effect on the translation of TOP mRNAs *in vivo*. *Nucleic Acids Res.* **28**, 2927–2934 (2000).
17. Ørom, U. A., Nielsen, F. C. & Lund, A. H. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol. Cell* **30**, 460–471 (2008).
18. Evdokimova, V. *et al.* Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. *Cancer Cell* **15**, 402–415 (2009).
19. Lahat, G. *et al.* Vimentin is a novel anti-cancer therapeutic target; insights from *in vitro* and *in vivo* mice xenograft studies. *PLoS ONE* **5**, e10105 (2010).
20. Hofer, M. D. *et al.* The role of metastasis-associated protein 1 in prostate cancer progression. *Cancer Res.* **64**, 825–829 (2004).
21. Yoo, Y. G., Kong, G. & Lee, M. O. Metastasis-associated protein 1 enhances stability of hypoxia-inducible factor-1 $\alpha$  protein by recruiting histone deacetylase 1. *EMBO J.* **25**, 1231–1241 (2006).
22. Liu, C. *et al.* The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nature Med.* **17**, 211–215 (2011).
23. Okuzumi, T. *et al.* Inhibitor hijacking of Akt activation. *Nature Chem. Biol.* **5**, 484–493 (2009).
24. Dowling, R. J. *et al.* mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328**, 1172–1176 (2010).
25. Jacinto, E. *et al.* SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates Akt phosphorylation and substrate specificity. *Cell* **127**, 125–137 (2006).
26. Wang, X. *et al.* A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature* **461**, 495–500 (2009).
27. Mulholland, D. J. *et al.* Lin<sup>−</sup> Sca-1<sup>+</sup> CD49<sup>high</sup> stem/progenitors are tumor-initiating cells in the Pten-null prostate cancer model. *Cancer Res.* **69**, 8555–8562 (2009).
28. Wang, S. *et al.* Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer. *Cancer Cell* **4**, 209–221 (2003).
29. Sutherland, B. W. *et al.* Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. *Oncogene* **24**, 4281–4292 (2005).
30. Leong, K. G., Wang, B. E., Johnson, L. & Gao, W. Q. Generation of a prostate from a single adult stem cell. *Nature* **456**, 804–818 (2008).
31. Lang, S. H. *et al.* Enhanced expression of vimentin in motile prostate cell lines and in poorly differentiated and metastatic prostate carcinoma. *Prostate* **52**, 253–263 (2002).
32. Helfand, B. T. *et al.* Vimentin organization modulates the formation of lamellipodia. *Mol. Biol. Cell* **22**, 1274–1289 (2011).
33. Amato, R. J., Jac, J., Mohammad, T. & Saxena, S. Pilot study of rapamycin in patients with hormone-refractory prostate cancer. *Clin. Genitourin. Cancer* **6**, 97–102 (2008).
34. George, D. J. *et al.* A phase II study of RAD001 in men with hormone refractory metastatic prostate cancer (HRPC). *Am. Soc. Clin. Oncol. Genitourin. Cancers Symp.* Abstract 181 (2008).
35. Pontes, J. E., Wajsbman, Z., Huben, R. P., Wolf, R. M. & Englander, L. S. Prognostic factors in localized prostatic carcinoma. *J. Urol.* **134**, 1137–1139 (1985).
36. Zhou, P. *et al.* Predictors of prostate cancer-specific mortality after radical prostatectomy or radiation therapy. *J. Clin. Oncol.* **23**, 6992–6998 (2005).
37. Grolleau, A. *et al.* Global and specific translational control by rapamycin in T cells uncovered by microarrays and proteomics. *J. Biol. Chem.* **277**, 22175–22184 (2002).
38. Ruggero, D. R. *et al.* The translation factor eIF-4F promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature Med.* **10**, 484–486 (2004).
39. Willett, M., Brocard, M., Davide, A. & Morley, S. J. Translation initiation factors and active sites of protein synthesis co-localize at the leading edge of migrating fibroblasts. *Biochem. J.* **438**, 217–227 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Barna for critical discussion and reading of this manuscript; T. Wilson for support and advice; T. Sanders and E. Lagostera-Martin for technical support with confocal microscopy; L. Li, E. Ulm, L. Kessler, J. Kucharski and L. Darjania for technical support for the discovery and development of INK128. J. Kurhanewicz and R. Bok of the Surbeck Institute for Advanced Imaging for technical support and MRI images; N. Sonenberg for providing the 4EBP1/2 double knockout mouse embryonic fibroblasts; J. M. Shen for support; and K. Tong for editing the manuscript. A.C.H. is supported in part by the American Cancer Society (119084-PF-10-233-01-TBE), and is a Prostate Cancer Foundation Young Investigator, and a recipient of the DOD Prostate Cancer Training Award. This work is supported by NIH R01 CA154916 (D.R.), NIH R01 CA140456 (D.R.) and the Phi Beta Psi Sorority (D.R.). D.R. is a Leukemia & Lymphoma Society Scholar.

**Author Contributions** A.C.H. and D.R. conceived the experiments. A.C.H., M.P.E., M.R.J., A.S., E.Y.S., C.R.S., C.C. and S.W. performed the experiments, *Pten*<sup>−/L</sup> preclinical trials, and collected the data. N.T.I. and J.S.W. contributed to ribosomal profiling data analysis. M.J.B. provided pathology support. Y.L., P.R., M.M., S.W., K.J., M.E.F., K.M.S. and C.R. developed and/or supported development of INK128, conducted pharmacokinetic, pharmacodynamic and preclinical studies. A.C.H. and D.R. analysed the data and wrote the manuscript. All authors discussed results and edited the manuscript.

**Author Information** Small-RNA sequencing data were deposited in the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE35469. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.R. ([davide.ruggero@ucsf.edu](mailto:davide.ruggero@ucsf.edu)) or C.R. ([christian@intellikine.com](mailto:christian@intellikine.com)).

## METHODS

**Mice.** *Pten*<sup>loxP/loxP</sup> and *Pb-cre* mice were obtained from Jackson Laboratories and Mouse Models of Human Cancers Consortium (MMHCC), respectively, and maintained in the C57BL/6 background. Mice were maintained under specific pathogen-free conditions, and experiments were performed in compliance with institutional guidelines as approved by the Institutional Animal Care and Use Committee of UCSF.

**Cell culture and reagents.** Human cell lines were obtained from the ATCC and maintained in the appropriate medium with supplements as suggested by ATCC. Wild-type, *mSin1*<sup>-/-</sup> (provided by B. Su), and *4EBP1/4EBP2* double knockout MEFs (provided by N. Sonenberg) were cultured as previously described<sup>24,25</sup>. SMARTvector 2.0 (Thermo Scientific) lentiviral shRNA constructs were used to knock down PTEN (SH-003023-02-10). For generation of GFP-labelled PC3 cells, SMARTvector 2.0 lentiviral empty vector control particles that contain TurboGFP (S-004000-01) were used. Control (D-001810-01), *YB1* (L-010213), *MTA1* (L-004127), *CD44* (L-009999), vimentin (L-003551), rictor (LL-016984), *4EBP1* (L-003005) and *4EBP2* (L-018671) pooled siRNAs were purchased from Thermo Scientific. Intellikine provided INK128 and PP242, which were used at 200 nM and 2.5  $\mu$ M in cell-based assays unless otherwise specified. RAD001 was obtained from LC Laboratories. DG-2 was provided by K. Shokat and used at 20  $\mu$ M in cell-based assays. Rapamycin was purchased from Calbiochem and used at 50 nM in cell-based assays. Doxycycline (Sigma) was used at 1  $\mu$ g ml<sup>-1</sup> in *4EBP1*<sup>M</sup> induction assays. Lipofectamine 2000 (Invitrogen) was used to transfect cancer cell lines with siRNA. Amaxa Cell Line Nucleofector Kit R (Lonza) was used to electroporate BPH-1 cells with over expression vectors. The *4EBP1*<sup>M</sup> has been previously described<sup>13</sup>.

**Plasmids.** pcDNA3-HA-YB1 was provided by V. Evdokimova. pCMV6-Myk-DDK-MTA1 was purchased from Origene. pGL3-Promoter was purchased from Promega. To clone the 5' UTR of *YB1* into pGL3-Promoter, the entire 5' UTR sequence of *YB1* was amplified from PC3 cDNA. PCR fragments were digested with HindIII and NcoI and ligated into the corresponding sites of pGL3-Promoter. The PRTE sequence at position +20–34 in the *YB1* 5' UTR (UCSC kgID uc001chs.2) was mutated using the QuikChange Site-Directed Mutagenesis Kit following the manufacturer's protocol (Stratagene).

**Ribosome profiling.** PC3 cells were treated with rapamycin (50 nM; Calbiochem) or PP242 (2.5  $\mu$ M; Intellikine) for 3 h. Cells were subsequently treated with cycloheximide (100  $\mu$ g ml<sup>-1</sup>; Sigma) and detergent lysis was performed in the dish. The lysate was treated with DNase and clarified, and a sample was taken for RNA-seq analysis. Lysates were subjected to ribosome foot printing by nuclease treatment. Ribosome-protected fragments were purified, and deep sequencing libraries were generated from these fragments, as well as from poly(A) mRNA purified from non-nuclease-treated lysates. These libraries were analysed by sequencing on an Illumina GAI.

Each sequencing run resulted in approximately 20–25 million raw reads per sample, of which 5–12 million unique reads were used for subsequent analysis. Ribosome footprint and RNA-seq sequencing reads were aligned against a library of transcripts from the UCSC Known Genes database GRCh37/hg19. The first 25 nucleotides of each read were aligned using Bowtie and this initial alignment was then extended to encompass the full fragment-derived portion of the sequencing read while excluding the linker sequence. Read density profiles were then constructed for the canonical transcript of each gene, using only reads with 0 or 1 total mismatches between the read sequence and the reference sequence, comprised of the transcript fragment followed by the linker sequence. Footprint reads were assigned to an A site nucleotide at position +15 to +17 of the alignment, based on the total fragment length; mRNA reads were assigned to the first nucleotide of the alignment. The average read density per codon was then computed for the coding sequence of each transcript, excluding the first 15 and last 5 codons, which can display atypical ribosome accumulation.

Average read density was used as a measure of mRNA abundance (RNA-seq reads) and of protein synthesis (ribosome profiling reads). For most analyses, genes were filtered to require at least 256 reads in the relevant RNA-seq samples. Translational efficiency was computed as the ratio of ribosome footprint read density to RNA-seq read density, scaled to normalize the translational efficiency of the median gene to 1.0 after excluding regulated genes (log<sub>2</sub> fold-change  $\pm$  1.5 after normalizing for the all-gene median). Changes in protein synthesis, mRNA abundance and translational efficiency were similarly computed as the ratio of read densities between different samples, normalized to give the median gene a ratio of 1.0. This normalization corrects for differences in the absolute number of sequencing reads obtained for different libraries. 3,977 (replicate 1), and 5,333 (replicate 2) unique mRNAs passed a preset read threshold of 256 reads for single-gene quantification for all treatment conditions.

**Western blot analysis.** Western blot analysis was performed as previously described<sup>13</sup> with antibodies specific to phospho-AKT<sup>S473</sup> (Cell Signaling), AKT

(Cell Signaling), phospho-p70S6K<sup>T389</sup> (Cell Signaling), phospho-rpS6<sup>S240/244</sup> (Cell Signaling), rpS6 (Cell Signaling), phospho-4EBP1<sup>T37/46</sup> (Cell Signaling), 4EBP1 (Cell Signaling), 4EBP2 (Cell Signaling), YB1 (Cell Signaling), CD44 (Cell Signaling), LEF1 (Cell Signaling), PTEN (Cell Signaling), eEF2 (Cell Signaling), GAPDH (Cell Signaling), vimentin (BD Biosciences), eIF4E (BD Biosciences), Flag (Sigma),  $\beta$ -actin (Sigma), MTA1 (Santa Cruz Biotechnology), Twist (Santa Cruz Biotechnology), rpL28 (Santa Cruz Biotechnology), HA (Covance) and rictor (Bethyl Laboratory).

**qPCR analysis.** RNA was isolated using the manufacturer's protocol for RNA extraction with TRIzol Reagent (Invitrogen) using the Pure Link RNA mini kit (Invitrogen). RNA was Dnase-treated with Pure Link Dnase (Invitrogen). Dnase-treated RNA was transcribed to cDNA with SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen), and 1  $\mu$ l of cDNA was used to run a SYBR green detection qPCR assay (SYBR Green Supermix and MyiQ2, Biorad). Primers were used at 200 nM.

**5' UTR analysis.** 5' UTRs of the 144 downregulated mTOR target genes were obtained using the known gene ID from the UCSC Genome Browser (GRCh37/hg19). Target versus non-target mRNAs were compared for 5' UTR length, %G+C content and Gibbs free energy by the Wilcoxon two-sided test. Multiple  $E_m$  (expectation maximization) for Motif Elicitation (MEME) and Find Individual Motif Occurrences (FIMO) was used to derive the PRTE and determine its enrichment in the 144 mTOR-sensitive genes compared a background list of 3,000 genes. The Database of Transcriptional Start Sites (DBTSS Release 8.0) was used to identify putative 5' TOP genes and putative transcription start sites in the 144 mTOR target genes.

**Luciferase assay.** PC3 *4EBP1*<sup>M</sup> cells were treated with 1  $\mu$ g ml<sup>-1</sup> doxycycline (Sigma) for 24 h. Cells were transfected with various pGL3-Promoter constructs using lipofectamine 2000 (Invitrogen). After 24 h, cells were collected. 20% of the cells were aliquoted for RNA isolation. The remaining cells were used for the luciferase assay per the manufacturer's protocol (Promega). Samples were measured for luciferase activity on a Glomax 96-well plate luminometer (Promega). Firefly luciferase activity was normalized to luciferase mRNA expression levels.

**Kinase assays.** mTOR activity was assayed using LanthaScreen Kinase kit reagents (Invitrogen) according to the manufacturer's protocol. PI(3)K  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  activity were assayed using the PI(3)K HTRF assay kit (Millipore) according to the manufacturer's protocol. The concentration of INK128 necessary to achieve inhibition of enzyme activity by 50% (IC<sub>50</sub>) was calculated using concentrations ranging from 20  $\mu$ M to 0.1 nM (12-point curve). IC<sub>50</sub> values were determined using a nonlinear regression model (GraphPad Prism 5).

**Cell proliferation assay.** PC3 cells were treated with the appropriate drug for 48 h, and proliferation was measured using CellTiter-Glo Luminescent reagent (Promega) per the manufacturer's protocol. The concentration of INK128 necessary to achieve inhibition of cell growth by 50% (IC<sub>50</sub>) was calculated using concentrations ranging from 20.0  $\mu$ M to 0.1 nM (12-point curve).

**Mouse xenograft study.** Nude mice were inoculated subcutaneously in the right subscapular region with  $5 \times 10^6$  MDA-MB-361 cells. After tumours reached a size of 150–200 mm<sup>3</sup>, mice were randomly assigned into vehicle control or treatment groups. INK128 was formulated in 5% polyvinylpyrrolidone, 15% NMP, 80% water and administered by oral gavage at 0.3 mg kg<sup>-1</sup> and 1 mg kg<sup>-1</sup> daily.

**Pharmacokinetic analysis.** The area under the plasma drug concentration versus time curves, AUC<sub>(0–t<sub>last</sub>)</sub> and AUC<sub>(0–inf)</sub>, were calculated from concentration data using the linear trapezoidal rule. The terminal  $t_{1/2}$  in plasma was calculated from the elimination rate constant ( $\lambda_z$ ), estimated as the slope of the log-linear terminal portion of the plasma concentration versus time curve, by linear regression analysis. The bioavailability ( $F$ ) was calculated using  $F = (AUC_{(0–last),po} D_{i.v.}) / (AUC_{(0–last),iv} D_{p.o.}) \times 100\%$ , where  $D_{i.v.}$  and  $D_{p.o.}$  are intravenous and oral doses, respectively.  $C_{max}$  was a highest drug concentration in plasma after oral administration.  $T_{max}$  was the time at which  $C_{max}$  is observed after extravascular administration of drug.  $T_{last}$  was the last time point a quantifiable drug concentration can be measured.

**Metabolic stability assay.** *In vitro* metabolic stability of INK128 was evaluated after incubation with liver microsomes or liver S9 fractions from various species in the presence of NADPH. The half-life of INK128 was estimated by log linear regression analysis.

**CYP assay.** INK128 inhibition of CYP450 isoforms in human liver microsomes was determined with isoform-specific substrates at concentrations approximately equal to the concentration at which the rate of the reaction is half-maximal ( $K_m$ ) for the individual isoforms: CYP1A2, CYP2C8, CYP2C9, CYP2C19, CYP2D6 and CYP3A4.

**Pharmaceutical property assays.** The percentage of protein binding of INK128 was determined in mouse, rat, dog, monkey and human plasma at CEREP. The IC<sub>50</sub> for the inhibitory effect of INK128 on hERG potassium channel was determined at



CEREP. A Bacterial Reverse Mutation Assay (Ames test) was conducted at BioReliance.

**Polysome analysis.** PC3 cells were treated for 3 h with either DMSO or INK128 (100 nM). Cells were re-suspended in PBS containing 100  $\mu\text{g ml}^{-1}$  cycloheximide (Sigma) and incubated on ice for 10 min. Cells were centrifuged at 300g for 5 min at 4 °C and lysed in 10 mM Tris-HCl pH 8, 140 mM NaCl, 5 mM  $\text{MgCl}_2$ , 640 U  $\text{ml}^{-1}$  Rnase, 0.05% NP-40, 250  $\mu\text{g ml}^{-1}$  cycloheximide, 20 mM DTT and protease inhibitors. Samples were incubated for 20 min on ice then centrifuged once for 5 min at 3,300g and once for 5 min at 9,300g, isolating the supernatant after each centrifugation. Lysates were loaded onto 10–50% sucrose gradients containing 0.1 mg  $\text{ml}^{-1}$  heparin and 2 mM DTT and centrifuged at 37,000 r.p.m. for 2.5 h at 4 °C. The sample was subsequently fractionated on a gradient fractionation system (ISCO). RNA was extracted from all fractions and run on a TBE-agarose gel to visualize 18S and 28S rRNA. Fractions 7–13 were found to correspond to the polysome fractions and were used for further qPCR analysis.

**[ $^{35}\text{S}$ ] metabolic labelling.** PC3 or PC3 4EBP1<sup>M</sup> cells with or without indicated treatment were incubated with 30  $\mu\text{Ci}$  of [ $^{35}\text{S}$ ]-methionine for 1 h after pre-incubation in methionine-free DMEM (Invitrogen). Cells were prepared using a standard protein lysate protocol, resolved on a 10% SDS polyacrylamide gel and transferred onto a PVDF membrane (Biorad). The membrane was exposed to autoradiography film (Denville) for 24 h and developed.

**Cell cycle analysis.** Appropriately treated PC3, BPH-1, or PC3-4EBP1<sup>M</sup> cells were fixed in 70% ethanol overnight at –20 °C. Cells were subsequently washed with PBS and treated with RNase (Roche) for 30 min. After this incubation, the cells were permeabilized and treated with 50  $\mu\text{g ml}^{-1}$  propidium iodide (Sigma) in a solution of 0.1% Tween, 0.1% sodium citrate. Cell cycle data was acquired using a BD FACS Caliber (BD Biosciences) and analysed with FlowJo (v.9.1).

**Apoptosis analysis.** Appropriately treated LNCaP and A498 cells were labelled with Annexin V-FITC (BD Biosciences) and propidium iodide (Sigma) following the manufacturer's instructions. PI/Annexin data was acquired using a BD FACS Caliber (BD Biosciences) and analysed with FlowJo (v.9.1).

**Matrigel invasion assay.** BioCoat Matrigel Invasion Chambers (modified Boyden Chamber Assay; BD Biosciences) were used according to the manufacturer's instructions.

**Real-time imaging of cell migration.** Real-time imaging of GFP-labelled PC3 cells was performed in poly-D-lysine-coated chamber cover glass slides (Lab-Tek). PC3 GFP cells were plated and allowed to adhere for 24 h. Wells were wounded with a P200 pipette tip. The chamber slides were imaged with an IX81 Olympus wide-field fluorescence microscope equipped with a CO<sub>2</sub> and temperature controlled chamber and time-lapse tracking system. Images from DIC and GFP channels were taken every 2 min and processed using ImageJ (<http://rsb.info.nih.gov/ij/>) and analysed for cell migration with Manual Tracking (<http://rsbweb.nih.gov/ij/plugins/track/track.html>), using local maximum centring correction to maintain a centroid xy coordinate for each cell per frame over time. Tracking data was subsequently processed with the Chemotaxis and Migration tool from ibidi ([http://www.ibidi.de/applications/ap\\_chemo.html](http://www.ibidi.de/applications/ap_chemo.html)) to create xy coordinate plots, velocity and distance measurements.

**Snail1 immunocytochemistry.** Appropriately transfected or treated PC3 cells were plated on a poly-L-lysine-coated chamber slide (Lab-Tek) and cultured for 48 h. Cells were fixed with 4% paraformaldehyde (EMS), rinsed with PBS and permeabilized with 0.1% Triton X-100. The samples were blocked in 5% goat serum and then incubated with anti-Snail1 antibody (Cell Signaling) in 5% goat serum for 2 h at room temperature. Cells were washed with PBS and incubated with Alexa 594 anti-mouse antibody (Invitrogen) and DAPI (Invitrogen) for 2 h at room temperature. Specimens were again washed with PBS and subsequently mounted with Aqua Poly/Mount (Polysciences). Image capture and quantification were completed as described below (see Immunofluorescence).

**Cap-binding assay.** PC3 4EBP1<sup>M</sup> cells were induced with doxycycline (1  $\mu\text{g ml}^{-1}$ , Sigma) for 48 h, then collected and lysed in buffer A (10 mM Tris-HCl pH 7.6, 150 mM KCl, 4 mM  $\text{MgCl}_2$ , 1 mM DTT, 1 mM EDTA, and protease inhibitors, supplemented with 1% NP-40). Cell lysates were incubated overnight at 4 °C with 50  $\mu\text{l}$  of the mRNA cap analogue m<sup>7</sup>GTP-sepharose (GE Healthcare) in buffer A. The beads were washed with buffer A supplemented with 0.5% NP-40. Protein complexes were dissociated using 1× sample buffer, and resolved by SDS-PAGE and western blotted with the appropriate antibodies.

**Pharmacological treatment of *Pten*<sup>L/L</sup> mice and MRI imaging.** Nine- and twelve-month-old *Pten*<sup>L/L</sup> mice were gavaged daily with either vehicle (see mouse xenograft study), RAD001 (10 mg  $\text{kg}^{-1}$ ; LC Laboratories), or INK128 (1 mg  $\text{kg}^{-1}$ ; Intellikine) for the indicated times. Weight measurements were taken every 3 days to monitor for toxicity. For the 28-day study, mice were imaged via MRI at day 0 and day 28 in a 14-T GE MR scanner (GE Healthcare).

**Prostate tissue processing.** Whole mouse prostates were removed from wild-type and *Pten*<sup>L/L</sup> mice, microdissected, and frozen in liquid nitrogen. Frozen

tissues were subsequently manually disassociated using a biopulverizer (Biospec) and additionally processed for protein and mRNA analysis as described above.

**Immunofluorescence.** Prostates and lymph nodes were dissected from mice within 2 h of the indicated treatment and fixed in 10% formalin overnight at 4 °C. Tissues were subsequently dehydrated in ethanol (Sigma) at room temperature, mounted into paraffin blocks, and sectioned at 5  $\mu\text{m}$ . Specimens were de-paraffinized and rehydrated using CitriSolv (Fisher) followed by serial ethanol washes. Antigen unmasking was performed on each section using Citrate pH 6 (Vector Labs) in a pressure cooker at 125 °C for 10–30 min. Sections were washed in distilled water followed by TBS washes. The sections were then incubated in 5% goat serum, 1% BSA in TBS for 1 h at room temperature. Various primary antibodies were used including those specific for keratin 5 (Covance), cytokeratin 8 (Abcam and Covance), YB1 (Abcam), vimentin (Abcam), MTA1 (Cell signaling), CD44 (BD Pharmingen) and the androgen receptor (Epitomics), which were diluted 1:50–1:500 in blocking solution and incubated on sections overnight at 4 °C. Specimens were then washed in TBS and incubated with the appropriate Alexa 488 and 594 labelled secondary (Invitrogen) at 1:500 for 2 h at room temperature with the exception of YB1 which was incubated with biotinylated anti-rabbit secondary (Vector) followed by incubation with Alexa 594 labelled Streptavidin (Invitrogen). A final set of washes in TBS was completed at room temperature followed by mounting with DAPI Hardset Mounting Medium (Vector Lab). A Zeiss Spinning Disc confocal (Zeiss, CSU-X1) was used to image the sections at 40×–100×. Individual prostate cells were quantified for mean fluorescence intensity (m.f.i.) using the Axiovision (Zeiss, Release 4.8) densitometric tool.

**Lymph node metastasis measurements.** Mouse lymph nodes were processed as described above and stained for CK8 and androgen receptor. Lymph nodes were imaged using a Zeiss AX10 microscope. Metastases were identified and areas were measured using the Axiovision (Zeiss, Release 4.8) measurement tool.

**Semi-quantitative RT-PCR.** Whole prostates were removed from wild-type and *Pten*<sup>L/L</sup> mice, microdissected, dissociated into single-cell suspension, and stained for epithelial cell markers as previously described<sup>40</sup> using fluorescence-conjugated antibodies for CD49f, Sca-1, CD31, CD45 and Ter119 (BD Biosciences). Luminal epithelial cells were sorted as previously described<sup>41</sup> using a FACS Aria (BD Biosciences). Cell pellets were re-suspended in 500  $\mu\text{l}$  TRIzol Reagent and RNA was isolated and transcribed into cDNA as described above. Semi-quantitative PCR analysis was performed using oligonucleotides for vimentin and  $\beta$ -actin at 200 nM in a 25  $\mu\text{l}$  reaction with 12.5  $\mu\text{l}$  GoTaq (Promega) for 32 and 33 cycles respectively, which were within the linear range (Supplementary Fig. 25f).

**Immunohistochemistry.** Immunohistochemistry was performed as described above (see immunofluorescence section) with the exception that immediately after antigen presentation and TBS washes, specimens were incubated in 3% hydrogen peroxide in TBS followed by TBS washes. The following primary antibodies were used: phospho-AKT<sup>S473</sup> (Cell Signaling), phospho-rpS6<sup>S240/244</sup> (Cell Signaling), phospho-4EBP1<sup>T37/46</sup> (Cell Signaling), phospho-histone H3 (Upstate), and cleaved caspase 3 (Cell Signaling). This was followed by TBS washes and incubation with the appropriate biotinylated secondary antibody (Vector Lab) for 30 min at room temperature. An ABC-HRP Kit (Vector Lab) was used to amplify the signal, followed by a brief incubation in hydrogen peroxide. The protein of interest was detected using DAB (Sigma). Specimens were counterstained with haematoxylin (Thermo Scientific), dehydrated with Citrisolv (Fisher), and mounted with Cytoseal XYZ (Vector Lab).

**Haematoxylin and eosin staining.** Paraffin-embedded prostate specimens were deparaffinized and rehydrated as described above (see immunofluorescence section), stained with haematoxylin (Thermo Scientific), and washed with water. This was followed by a brief incubation in differentiation RTU (VWR) and two washes with water followed by two 70% ethanol washes. The samples were then stained with eosin (Thermo Scientific) and dehydrated with ethanol followed by CitriSolv (Fisher). Slides were mounted with Cytoseal XYZ (Richard Allan Scientific).

**Oligonucleotides.** YB1 5' UTR cloning and site-directed mutagenesis oligonucleotides are as follows. YB1 5' UTR cloning: forward 5'-GCTACAGCTTGG GCTTATCCCGCCT-3', reverse 5'-TCGATCCATGGGGTGTGCGGTGATGGT-3'; deletion (20–34): forward 5'-TGGGCTTATCCCGCTGTCTTCGATCGGTA GCGGGAGCG-3', reverse 5'-CGCTCCCGCTACCGATCGGACAGGACAGGCG GGATAAGCCCA-3'; transversion (20–34): forward 5'-TGGGCTTATCCCGC CTGTCCCGGTAAGAGCGATCTTCGATCGGTAGCGGGAGCG-3', reverse 5'-CGCTCCCGCTACCGATCGAAGATCGCTCTTACCGCGGACAGGCGGG ATAAGCCCA-3'.

Human qPCR oligonucleotides are as follows.  $\beta$ -actin forward 5'-GCAA AGACCTGTACGCCAAC-3', reverse 5'-AGTACTTGCCTCAGGAGGA-3'; CD44 forward 5'-CAACAACACAAATGGCTGGT-3', reverse 5'-CTGAGGT GTCTGTCTCTTTCATCT-3'; vimentin forward 5'-GGCCAGCTGTAAGT TGGTA-3', reverse 5'-GGAGCGAGAGTGGCAGAG-3'; Snail1 forward

5'-CACTATGCCGCTCTTTC-3', reverse 5'-GCTGGAAGGTAAACTCTG GATTAGA-3'; *YB1* forward 5'-TCGCCAAAGACAGCCTAGAGA-3', reverse 5'-TCTGCGTCGGTAATTGAAGTTG-3'; *MTA1* forward 5'-CAAAGTGGTG TGCTTCTACCG-3', reverse 5'-CGGCCTTATAGCAGACTGACA-3'; *PLAU* forward 5'-TTGCTCACCACAACGACATT-3', reverse 5'-GGCAGGCAGATG GTCTGTAT-3'; *FGFBP1* forward 5'-ACTGGATCCGTGTGCTCAG-3', reverse 5'-GAGCAGGGTGAGGCTACAGA-3'; *ARID5B* forward 5'-TGGACTCAACT TCAAAGACGTTTC-3', reverse 5'-ACGTTTCGTTTCTTCCTCGTC-3'; *CTGF* forward 5'-CTCCTGCAGGCTAGAGAAGC-3', reverse 5'-GATGCACTTTT TGCCCTTCTT-3'; *RND3* forward 5'-AAAAACTGCGCTGCTCCAT-3', reverse 5'-TCAAAACTGGCCGTGTAAATTC-3'; *KLF6* forward 5'-AAAGCTC CCACTTGAAAGCA-3', reverse 5'-CCTTCCCATGAGCATCTGTAA-3'; *BCL6* forward 5'-TTCCGCTACAAGGGCAAC-3', reverse 5'-TGCAACGATA GGGTTTCTCA-3'; *FOXA1* forward 5'-AGGGCTGGATGGTTGTATTG-3', reverse 5'-ACCGGACGAGGAGTAG-3'; *GDF15* forward 5'-CCGGATAC TCACGCCAGA-3', reverse 5'-AGAGATACGCAGGTGCAGGT-3'; *HBPI* forward 5'-GCTGGTGGTGTGTGCTG-3', reverse 5'-CATGTTATGGTGCT CTGACTGC-3'; *Twist1* forward 5'-CATCCTCACACCTCTGCATT-3', reverse 5'-TTCCTTTTCAGTGGCTGATTG-3'; *LEF1* forward 5'-CCTTGGTGAACGA GTCTGAAATC-3', reverse 5'-GAGGTTTGTGCTGTCTGTC-3'; *rpS19* forward 5'-GCTGGCCAAACATAAAGAGC-3', reverse 5'-CTGGGTCTGAC ACCGTTTCT-3'; 5S rRNA forward 5'-GCCCAGATCTCGTCTGATCT-3', reverse 5'-AGCCTACAGCACCCGGTATT-3'; firefly luciferase forward 5'-AATCAAAGAGGCGAACTGTG-3', reverse 5'-TTCGTCTTCGTCCCAGT AAG-3'.

Mouse qPCR oligonucleotides are as follows.  $\beta$ -actin forward 5'-CTAAGG CCAACCGTGAAAAG-3', reverse 5'-ACCAGAGGCATACAGGGACA-3'; *Yb1* forward 5'-GGGTTACAGACCACGATTCC-3', reverse 5'-GGCGATACC GACGTTGAG-3'; vimentin forward 5'-TCCAGCAGCTTCCTGTAGGT-3',

reverse 5'-CCCTCACCTGTGAAGTGGAT-3'; *Cd44* forward 5'-ACAGTACCT TACCCACCATG-3', reverse 5'-GGATGAATCCTCGGAATTAC-3'; *Mta1* forward 5'-AGTGCGCCTAATCCGTGGTG-3', reverse 5'-CTGAGGATGAG AGCAGCTTTTCG-3'.

siRNA/shRNA sequences are as follows. Control (D-001810-01) 5'-UGGU UUACAUGUCGACUAA-3'; vimentin (L-003551) 5'-UCACGAUGACCUUG AAUAA-3', 5'-GGAAAUGGCUCGUCACCUU-3', 5'-GAGGGAAACUAAU CUGGAU-3', 5'-UUAAGACGGUUGAAACUAG-3'; *YB1* (L-010213) 5'-CUG AGUAAAUGCCGGCUUA-3', 5'-CGACGCAGACGCCAGAAA-3', 5'-GUA AGGAACGGAUAUGGUU-3', 5'-GCGGAGGCAGCAAAUGUUA-3'; *MTA1* (L-004127) 5'-UCACGGACAUUCAGCAAGA-3', 5'-GGACCAAAACCGCAG UAACA-3', 5'-GCAUCUUGUUGGACAUUU-3', 5'-CCAGCAUCAUUGA GUACUA-3'; *CD44* (L-009999) 5'-GAAUUAUAAACUGCCGCUUU-3', 5'-CA AGUGGACUCAACGGAGA-3', 5'-CGAAGAAGGUGUGGGCAGA-3', 5'-GAUCAACAGUGGCAAUGGA-3'; *4EBP1* (L-003005) 5'-CUGAUGGAGU GUCGGAACU-3', 5'-CAUCUAUGACCGGAAAUUC-3', 5'-GCAAUAGCCC AGAAGAUAA-3', 5'-GAGAUGGACAUUUAAAGCA-3'; *4EBP2* (L-018671) 5'-GCAGCUACCUCAUGACUUA-3', 5'-GGAGGAACUCGAAUCAUUU- 3', 5'-GCAAUUCUCCCAUGGCUCA-3', 5'-UUGAACAACUUGAACAA UC-3'; rictor (LL-016984) 5'-GACACAAGCACUUCGAUUA-3', 5'-GAAGAU UUAUUGAGUCCUA-3', 5'-GCGAGCUGAUGUAGAAUUA-3', 5'-GGGA AUACAACUCCAAAUA-3'; *PTEN* SH-003023-01-10 5'-GCTAAGAGAGGT TTCCGA-3', SH-003023-02-10 5'-AGACTGATGTGTATACGTA-3'.

40. Lukacs, R. U., Goldstein, A. S., Lawson, D. A., Cheng, D. & Witte, O. N. Isolation, cultivation and characterization of adult murine prostate stem cells. *Nature Protocols* **5**, 702–713 (2010).
41. Lawson, D. A., Zong, Y., Memarzadeh, S., Xin, L., Huang, J. & Witte, O. N. Basal epithelial stem cells are efficient targets for prostate cancer initiation. *Proc. Natl Acad. Sci. USA* **107**, 2610–2615 (2010).



# Regulation of circadian behaviour and metabolism by synthetic REV-ERB agonists

Laura A. Solt<sup>1\*</sup>, Yongjun Wang<sup>1\*</sup>, Subhashis Banerjee<sup>1</sup>, Travis Hughes<sup>1</sup>, Douglas J. Kojetin<sup>1</sup>, Thomas Lundasen<sup>1</sup>, Youseung Shin<sup>2</sup>, Jin Liu<sup>1</sup>, Michael D. Cameron<sup>2</sup>, Romain Noel<sup>2</sup>, Seung-Hee Yoo<sup>3</sup>, Joseph S. Takahashi<sup>3</sup>, Andrew A. Butler<sup>4</sup>, Theodore M. Kamenecka<sup>2</sup> & Thomas P. Burris<sup>1,5</sup>

**Synchronizing rhythms of behaviour and metabolic processes is important for cardiovascular health and preventing metabolic diseases. The nuclear receptors REV-ERB- $\alpha$  and REV-ERB- $\beta$  have an integral role in regulating the expression of core clock proteins driving rhythms in activity and metabolism. Here we describe the identification of potent synthetic REV-ERB agonists with *in vivo* activity. Administration of synthetic REV-ERB ligands alters circadian behaviour and the circadian pattern of core clock gene expression in the hypothalamus of mice. The circadian pattern of expression of an array of metabolic genes in the liver, skeletal muscle and adipose tissue was also altered, resulting in increased energy expenditure. Treatment of diet-induced obese mice with a REV-ERB agonist decreased obesity by reducing fat mass and markedly improving dyslipidaemia and hyperglycaemia. These results indicate that synthetic REV-ERB ligands that pharmacologically target the circadian rhythm may be beneficial in the treatment of sleep disorders as well as metabolic diseases.**

In mammals, most if not all tissues display a self-sustaining circadian molecular pacemaker that is responsible for aligning rhythms in various physiological functions. The suprachiasmatic nucleus (SCN) of the hypothalamus functions as the master circadian pacemaker, synchronizing behavioural and physiological rhythms to the environmental light/dark cycle<sup>1</sup>. Optimal coordination of rhythms in metabolic processes with nutrient availability involves signals emanating from the SCN and hypothalamus, as well as autonomous inputs from nutrient sensors responding to metabolic flux and body temperature<sup>2</sup>.

The mammalian molecular clock is composed of a transcriptional feedback loop where heterodimers of the transcription factors BMAL1 (brain and muscle ARNT-like protein 1) and CLOCK (circadian locomotor output cycles kaput) or NPAS2 (neuronal PAS domain-containing protein 2) activate the transcription of the period (*Per1*, *Per2* and *Per3*) and cryptochrome (*Cry1* and *Cry2*) genes. Subsequently the PER and CRY proteins feedback to inhibit BMAL1–CLOCK activity, resulting in a rhythmic, circadian pattern of expression of these genes<sup>3</sup>. The REV-ERB nuclear receptors have an important role in feedback regulation of the circadian oscillator. Both *Bmal1* and *Clock* are direct REV-ERB target genes<sup>4,5</sup> and loss of REV-ERB- $\alpha$  alters circadian behaviour<sup>4</sup>. The physiological ligand for REV-ERB- $\alpha$  and REV-ERB- $\beta$  was recently identified as haem<sup>6,7</sup>, and based on observations that REV-ERB activity is regulated by a small molecule ligand, we and others have sought to identify and characterize synthetic ligands<sup>8–11</sup>. Here we describe the development of REV-ERB ligands that allowed for characterization of the effects of modulation of this receptor *in vivo*.

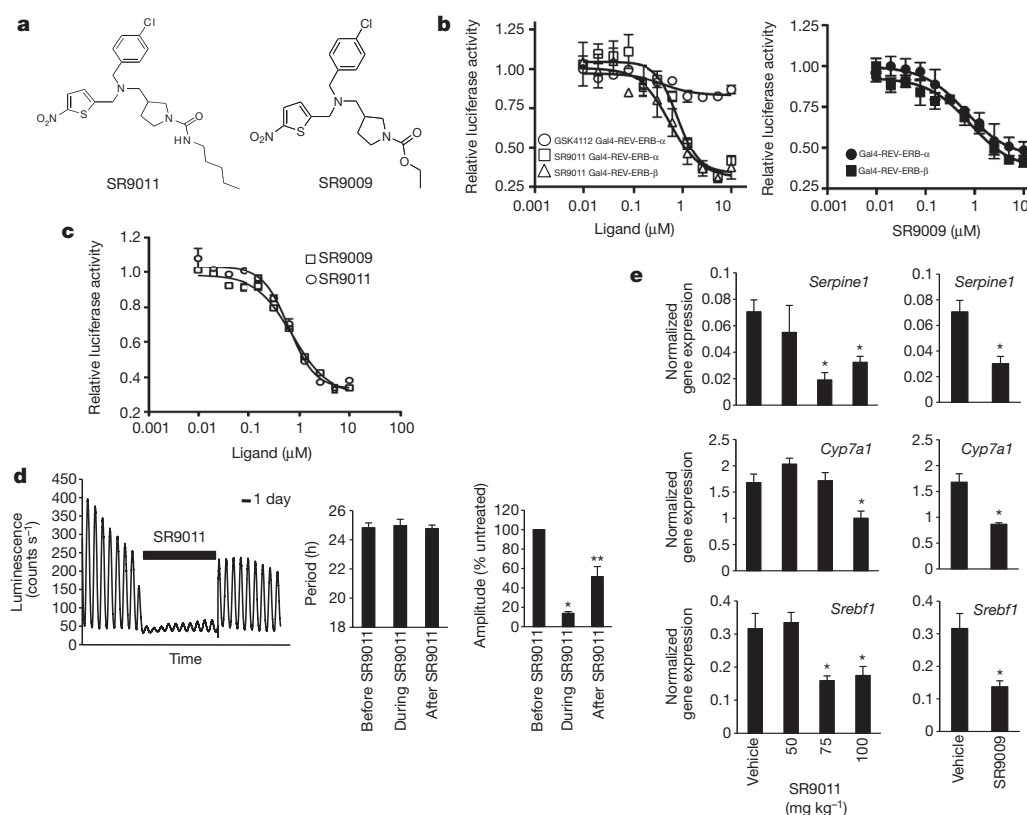
## Development of REV-ERB- $\alpha/\beta$ agonists

We developed two REV-ERB- $\alpha/\beta$  agonists with sufficient plasma/brain exposure to allow evaluation of their effects *in vivo*. Both SR9011 and

SR9009 (Fig. 1a and Supplementary Fig. 1) dose-dependently increased the REV-ERB-dependent repressor activity assessed in HEK293 cells expressing a chimaeric Gal4 DNA binding domain (DBD): REV-ERB ligand binding domain (LBD)  $\alpha$  or  $\beta$  and a Gal4-responsive luciferase reporter (Fig. 1b) (SR9009: REV-ERB- $\alpha$  half-maximum inhibitory concentration (IC<sub>50</sub>) = 670 nM, REV-ERB- $\beta$  IC<sub>50</sub> = 800 nM; SR9011: REV-ERB- $\alpha$  IC<sub>50</sub> = 790 nM, REV-ERB- $\beta$  IC<sub>50</sub> = 560 nM). The REV-ERB ligand GSK4112 (Supplementary Fig. 2), which exhibits no plasma exposure<sup>8,10</sup>, displays limited activity (Fig. 1b). Both SR9011 and SR9009 potently and efficaciously suppressed transcription in a co-transfection assay using full-length REV-ERB- $\alpha$  along with a luciferase reporter driven by the *Bmal1* promoter (Fig. 1c) (SR9009 IC<sub>50</sub> = 710 nM; SR9011 IC<sub>50</sub> = 620 nM). SR9011 and SR9009 suppressed the expression of *BMAL1* messenger RNA in HepG2 cells in a REV-ERB- $\alpha/\beta$ -dependent manner (Supplementary Fig. 3). Consistent with both compounds functioning as direct agonists of REV-ERB, we noted that the compounds increased the recruitment of the co-repressor-nuclear receptor box (CoNR) peptide fragment of nuclear receptor corepressor 1 (NCoR) using a biochemical assay (Supplementary Fig. 4)<sup>6</sup>. Direct binding of the SR9009 to REV-ERB- $\alpha$  was also confirmed using circular dichroism analysis (Supplementary Fig. 5) (dissociation (*K*<sub>d</sub>) = 800 nM). Neither compound exhibited activity at other nuclear receptors<sup>12,13</sup> (Supplementary Fig. 6). SR9011 also inhibited the activity of the SCN clock, with reversible inhibition of circadian oscillations in SCN explants cultured from the *Per2-luc* reporter mouse<sup>14</sup> (Fig. 1d). Treatment suppressed the amplitude of the oscillations but had no effect on the period (Fig. 1d). We observed similar effects in *Per2-luc* fibroblasts (Supplementary Fig. 7). The compounds displayed reasonable plasma exposure (Supplementary Fig. 8), thus, we examined the expression of REV-ERB-responsive genes in the

<sup>1</sup>Department of Molecular Therapeutics, The Scripps Research Institute, Jupiter, Florida 33458, USA. <sup>2</sup>Translational Research Institute, The Scripps Research Institute, Jupiter, Florida 33458, USA. <sup>3</sup>Howard Hughes Medical Institute and Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. <sup>4</sup>Department of Metabolism and Aging, The Scripps Research Institute, Jupiter, Florida 33458, USA. <sup>5</sup>Center for Diabetes and Metabolic Diseases, The Scripps Research Institute, Jupiter, Florida 33458, USA.

\*These authors contributed equally to this work.



**Figure 1 | SR9011 and SR9009 are synthetic REV-ERB agonists with activity *in vivo*.** **a**, Chemical structures of SR9011 and SR9009. **b**, Gal4-REV-ERB- $\alpha$  and Gal4-REV-ERB- $\beta$  co-transfection assays in HEK293 cells illustrating the activity of SR9011 and SR9009 and comparing the activity to GSK4112. **c**, Co-transfection assay in HEK293 cells with full-length REV-ERB- $\alpha$  and a luciferase reporter driven by the *Bmal1* promoter. **d**, Bioluminescence record from a

*Per2-luc* SCN treated with 5  $\mu$ M SR9011 as indicated by the bar. The right panels show the period and amplitude of the oscillations before, during and after treatment with SR9011. **e**, Expression of REV-ERB-responsive genes after treatment with various doses of SR9011 or 100 mg kg<sup>-1</sup> of SR9009 (i.p., twice per day (b.i.d.)) for 6 days. \* $P < 0.05$ ; \*\* $P < 0.05$  versus before SR9011 and during SR9011 treatment. Error bars indicate mean  $\pm$  s.e.m.;  $n = 6$ .

liver of mice treated with various doses of SR9011 for 6 days. The plasminogen activator inhibitor type 1 gene (*Serpine1*) is a REV-ERB target gene<sup>15</sup> and displayed dose-dependent suppression of expression in mice released into constant dark (D/D) conditions after 1 week of housing in wheel cages in a standard light/dark (L/D) setting. After 12 days in D/D conditions mice were injected with a single dose of SR9011, SR9009 or vehicle at circadian time 6 (CT6 (6 h after lights on); peak expression of *Rev-erb- $\alpha$*  (Supplementary Fig. 9)). Vehicle injection caused no disruption in circadian locomotor activity (Fig. 2a, top panels). However, administration of a single dose of either REV-ERB agonist resulted in loss of locomotor activity during the subject dark phase (Fig. 2a, bottom panels). Normal activity returned the next circadian cycle, consistent with clearance of the drugs in less than 24 h. This effect was not due to toxicity as the complete loss of locomotor activity was not observed in an identical experiment using L/D conditions (Fig. 2d). Additionally, mice treated with SR9011 did not display a decrease in strength (Supplementary Fig. 10a) and continued to move as detected in an open field assay (Supplementary Figs 10b, c). Furthermore, we observed no overt toxicity when we examined complete blood counts (Supplementary

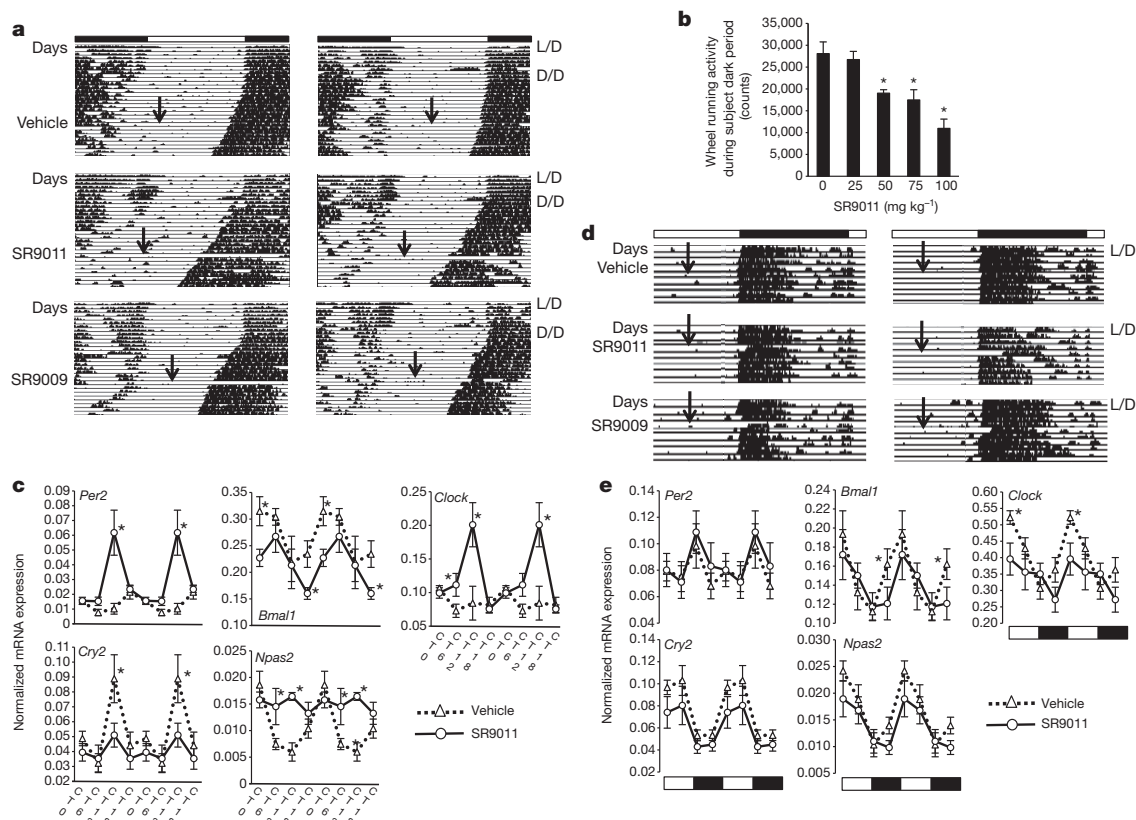
### REV-ERB agonists modulate circadian behaviour

On the basis of the effects of these compounds on SCN clock activity, we predicted that administration of these compounds would alter circadian behaviour. Circadian locomotor activity was examined in mice released into constant dark (D/D) conditions after 1 week of housing in wheel cages in a standard light/dark (L/D) setting. After 12 days in D/D conditions mice were injected with a single dose of SR9011, SR9009 or vehicle at circadian time 6 (CT6 (6 h after lights on); peak expression of *Rev-erb- $\alpha$*  (Supplementary Fig. 9)). Vehicle injection caused no disruption in circadian locomotor activity (Fig. 2a, top panels). However, administration of a single dose of either REV-ERB agonist resulted in loss of locomotor activity during the subject dark phase (Fig. 2a, bottom panels). Normal activity returned the next circadian cycle, consistent with clearance of the drugs in less than 24 h. This effect was not due to toxicity as the complete loss of locomotor activity was not observed in an identical experiment using L/D conditions (Fig. 2d). Additionally, mice treated with SR9011 did not display a decrease in strength (Supplementary Fig. 10a) and continued to move as detected in an open field assay (Supplementary Figs 10b, c). Furthermore, we observed no overt toxicity when we examined complete blood counts (Supplementary

Fig. 11). We observed that the SR9011-dependent decrease in wheel running behaviour in the mice under constant darkness conditions was dose-dependent (Fig. 2b) and that the potency (half-maximum effective dose (ED<sub>50</sub>) = 56 mg kg<sup>-1</sup>) was similar to the potency of SR9011-mediated suppression of a REV-ERB responsive gene, *Srebf1*, *in vivo* (ED<sub>50</sub> = 67 mg kg<sup>-1</sup>). *Tau* was not affected by treatment with either drug (data not shown), and the recovery after the drug to resume the normal rhythm is similar to the effect observed after removal of the drug from the SCN explants (Fig. 1d).

We next assessed the expression of core clock genes in hypothalami isolated from mice in D/D conditions. Mice were injected with a single dose of SR9011 or SR9009 at CT0 and hypothalami collected for expression analysis. We observed a range of effects on the pattern of expression of the core clock genes. The amplitude of *Per2* expression was enhanced whereas *Cry2* was suppressed (Fig. 2c). *Bmal1* expression was affected more subtly with a left shift in the phase of the circadian pattern (Fig. 2c). The circadian pattern of expression of *Npas2* was completely eliminated (Fig. 2c). The pattern of expression of *Clock* was also altered with SR9011 treatment, resulting in enhanced amplitude of the oscillation, but also altering the phase so that the *Clock* oscillation was in phase with the *Per2* oscillation (Fig. 2c). SR9009 treatment resulted in similar effects on gene expression (Supplementary Fig. 12). We also examined the effect of both REV-ERB ligands under L/D (12 h/12 h) conditions. Instead of complete loss of nocturnal locomotor activity, we noted a 1–3-h delay in the onset of nocturnal locomotor activity (Fig. 2d). Consistent with the more subtle effects on circadian behaviour, the effects of SR9011 and SR9009 on core clock gene expression in the hypothalamus were less severe than observed under constant darkness (Fig. 2e, Supplementary Fig. 13 and data not shown). Considerable differences





**Figure 2 | Synthetic REV-ERB ligands alter circadian behaviour and the pattern of expression of core clock genes.** **a**, Actograms illustrating the effect of single injections of vehicle, SR9011 (100 mg kg<sup>-1</sup>, i.p.) or SR9011 (100 mg kg<sup>-1</sup>, i.p.) on circadian behaviour. C57BL/6 mice were initially maintained on a 12 h/12 h L/D cycle and altered to constant darkness (D/D) after 7 days. After 12 days on D/D the mice were injected with vehicle or compound at CT6. **b**, Analysis of wheel running activity during the subject dark period after injection of SR9011 i.p. at CT6 in mice kept under constant darkness. **c**, Normalized expression levels of several core clock genes after

administration of SR9011 or vehicle under constant dark conditions. C57BL/6 mice were administered SR9011 (100 mg kg<sup>-1</sup>, i.p.) at CT0 on a day of constant darkness. Gene expression was determined and normalized to cyclophilin. Data are double plotted. **d**, Actograms illustrating the effect of single injections of vehicle, SR9011 or SR9009 in mice maintained under 12 h/12 h L/D conditions. **e**, Normalized expression levels of several core clock genes after administration of SR9011 or vehicle under L/D (12 h/12 h) conditions. Methods for **d** and **e** were otherwise identical to **a** and **c**. \**P* < 0.05. Error bars indicate mean ± s.e.m.; *n* = 6–10 mice.

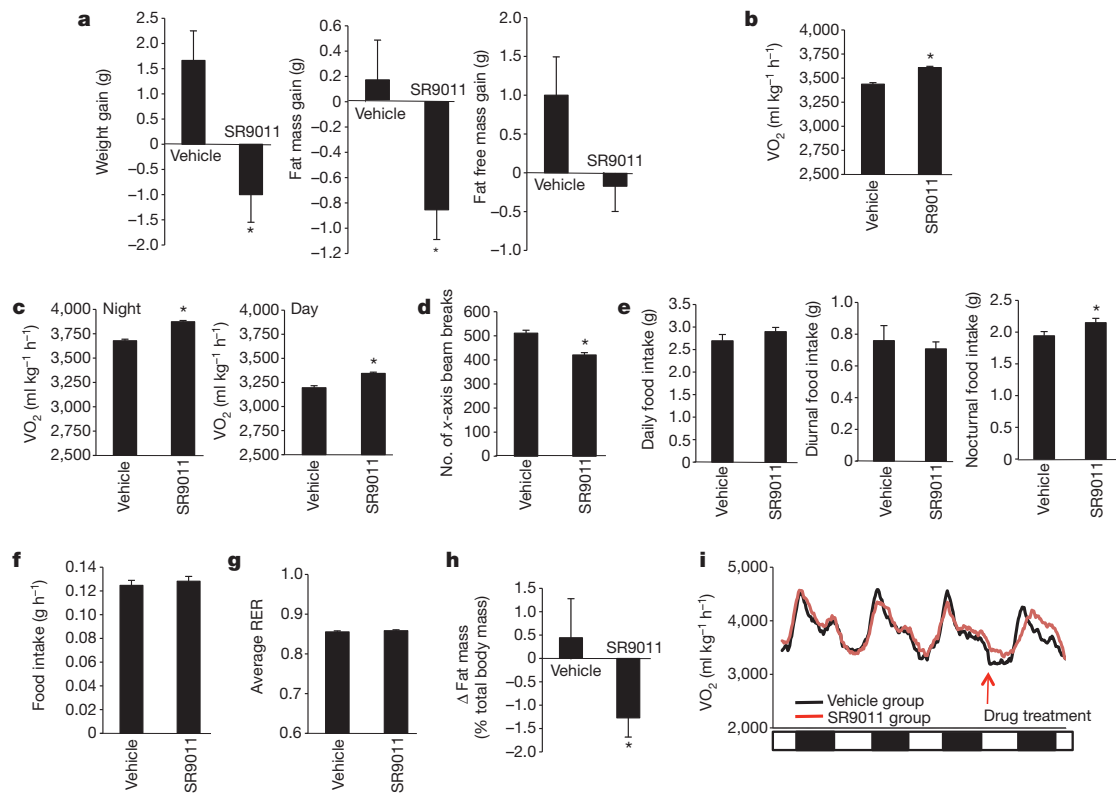
were noted in REV-ERB- $\alpha$  expression in the D/D versus L/D conditions in terms of their responsiveness to the compounds. In L/D conditions SR9011 had no effect on the circadian pattern of expression, whereas under D/D conditions SR9011 completely suppressed the circadian pattern of expression (Supplementary Fig. 9). These data indicate that light input into the circadian oscillator has a significant effect on the action of these drugs.

### REV-ERB- $\alpha/\beta$ agonists modulate metabolism in vivo

Genetic perturbation of the core clock genes leads to a range of metabolic phenotypes<sup>2,18</sup>. In addition, REV-ERB has been shown to directly regulate genes involved in lipid and glucose metabolism<sup>7,19,20</sup>. We observed clear metabolic effects when SR9011 was chronically administered to BALB/c mice. Mice displayed weight loss due to decreased fat mass (Fig. 3a); however, food intake was not affected (data not shown). Similar results were obtained in SR9009 treatment of C57BL/6 mice (Supplementary Fig. 14). We examined the metabolic effects of SR9011 in more detail using a comprehensive laboratory animal monitoring system (CLAMS). After acclimation, the animals were administered SR9011 twice per day for 10 days. A 5% increase in oxygen consumption (VO<sub>2</sub>) was observed, indicating increased energy expenditure (Fig. 3b). The increase in VO<sub>2</sub> was evident in the diurnal and nocturnal phases (Fig. 3c). The increases in VO<sub>2</sub> were not due to increased activity, as mice displayed a 15% decrease in movement (Fig. 3d). Treatment also had no effect on total daily food intake (Fig. 3e) or the rate of food intake (Fig. 3f), although there was a 10% increase noted in nocturnal food consumption in the SR9011

group. There was also no change in the respiratory exchange ratio (Fig. 3g). Consistent with increased metabolic rate, we observed a decrease in fat mass with SR9011 versus vehicle treatment (Fig. 3h). Treatment of mice housed in wheel cages in a L/D setting indicated a delayed onset of physical activity (Fig. 2c), and a similar 1–3 h delay in peak VO<sub>2</sub> was observed with administration of SR9011 (Fig. 3i). Given the association between the circadian rhythm and metabolic regulation, and to understand the potential mechanism underlying the alterations in metabolic rate, we examined the effect of the REV-ERB ligands on the circadian expression of various genes in the liver, muscle and adipose tissue. After a single injection of SR9011, we monitored the expression of clock genes in the liver over a 24-h period. When examining the effects of SR9011 treatment on core clock gene expression, the pattern of expression of *Per2* was altered, but expression of other genes such as *Bmal1* and *Npas2* was unaffected (Fig. 4a). Thus, SR9011 treatment results in alterations in the pattern of circadian expression of clock genes in both the hypothalamus (Fig. 2b, d and Supplementary Fig. 9) and liver, but there are clear distinctions in which genes are affected. These data suggest that the REV-ERB ligands differentially affect the central and peripheral clocks.

We also assessed the expression of an array of genes involved in metabolism in the liver in response to SR9011 treatment. The expression of lipogenic genes (*Srebf1*), fatty acid synthase (*Fasn*) and stearoyl-CoA-desaturase 1 (*Scd1*) were clearly altered with SR9011 treatment (Fig. 4b). Both *Srebf1* and *Scd1* expression were suppressed whereas the phase of *Fasn* expression was shifted (Fig. 4b). Expression of genes involved in cholesterol and bile acid metabolism were also



**Figure 3 | Activation of REV-ERB by SR9011 *in vivo* results in an increase in energy expenditure and weight loss.** **a**, Treatment of mice (BALB/c) with SR9011 results in weight loss and fat mass loss. Animals were dosed with SR9011 (100 mg kg<sup>-1</sup>, i.p., b.i.d.) for 12 days. **b**, Oxygen consumption (VO<sub>2</sub>) is increased in mice treated with SR9011. Results were obtained using CLAMS and C57BL/6 mice were dosed as described in **a** except that the duration of treatment was 10 days. **c**, Oxygen consumption (VO<sub>2</sub>) is increased during both the diurnal and nocturnal phases of C57BL/6 mice when they are treated with SR9011. Data obtained from the experiment described in **b** were analysed for time of day differences. **d**, Mice treated with SR9011 are less active in the

CLAMS as detected by the number of x-axis beam breaks. **e**, Total daily, diurnal and nocturnal food intake from the animals in the CLAMS study. **f**, The rate of food intake is not altered by SR9011 treatment. **g**, Respiratory exchange ratio (RER) is not altered by SR9011 treatment. **h**, After completion of the 10-day CLAMS experiment animals fat mass was assessed by DEXA. **i**, Results from a CLAMS experiment showing the diurnal increase in oxygen consumption before and immediately after administration of SR9011. Note the ~3-h delay in the diurnal peak in VO<sub>2</sub> after administration of SR9011. For all b.i.d. dosing animals were dosed at CT0 and CT12. \**P* < 0.05. Error bars indicate mean ± s.e.m.; *n* = 6–10 mice.

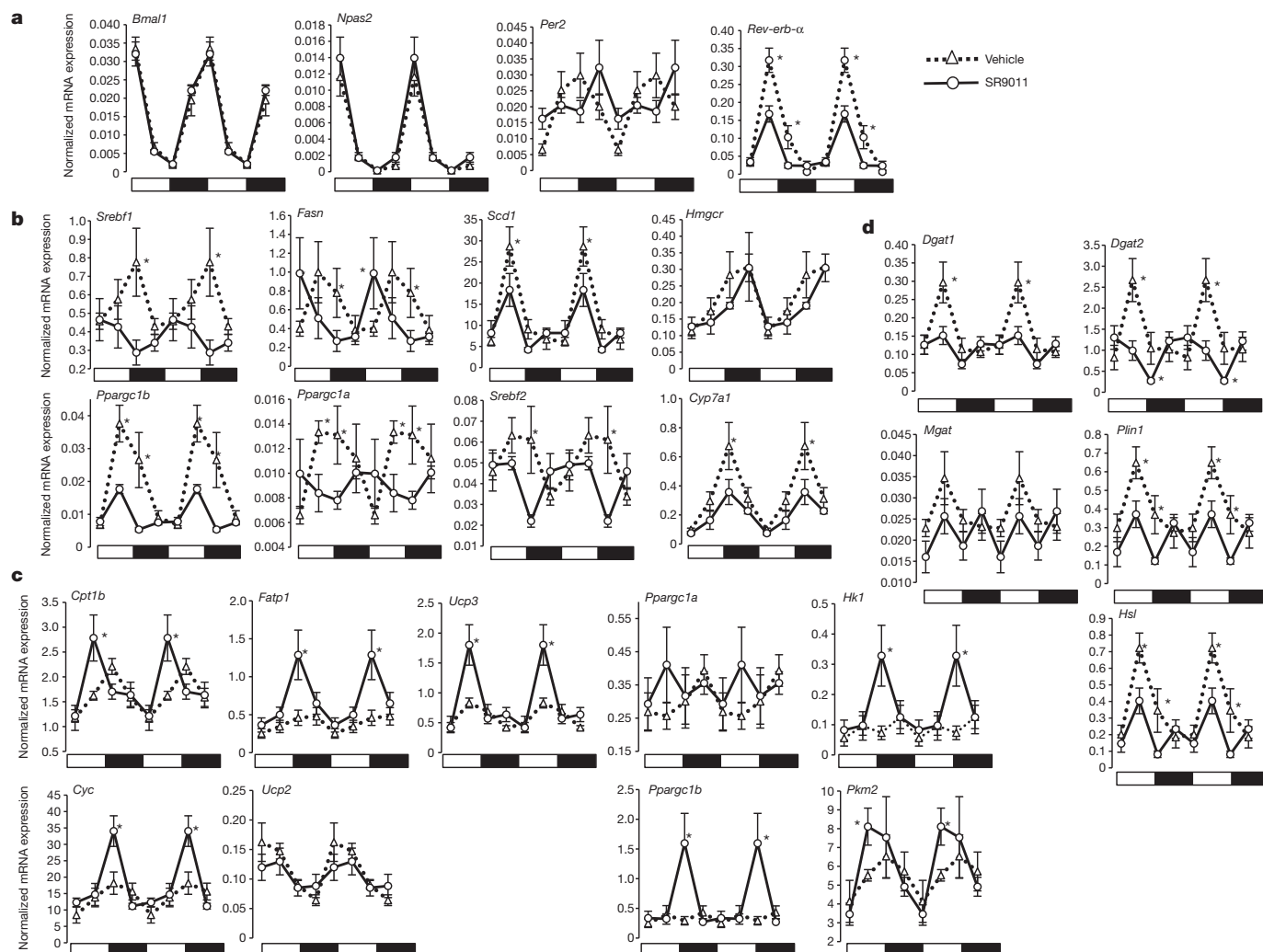
altered. Expression of *Srebf2* and *Cyp7a1* was decreased (Fig. 4b). 3-Hydroxy-3-methyl-glutaryl-CoA reductase (*Hmgcr*) was unaffected in this acute model (Fig. 4b). Peroxisome proliferator-activated receptor-γ coactivator 1α and 1β (*Ppargc1a* and *Ppargc1b*) both displayed a strong circadian pattern of expression that was suppressed with SR9011 treatment (Fig. 4b). Examination of gene expression in skeletal muscle revealed a potential mechanism for the increased metabolic rate that we observed in the CLAMS experiments. Expression levels of the genes that encode the rate-limiting enzyme for β-oxidation of fatty acids—carnitine palmitoyltransferase 1b (*Cpt1b*)—as well as fatty acid transport into the skeletal muscle—fatty acid transport protein 1 (*Fatp1*; also called *Slc27a1*)—were elevated (Fig. 4c). *Pppargc1b* expression was also elevated along with uncoupling protein 3 (*Ucp3*), consistent with altered fatty acid metabolism in skeletal muscle (Fig. 4c). The pattern of expression followed the expected diurnal increase in expression of genes involved in fatty acid oxidation, but the increases in expression were amplified (Fig. 4c). When we examined the expression of key enzymes in the glycolytic pathway (hexokinase (*Hk1*) and pyruvate kinase (*Pkm2*)) we noted an increase in both (Fig. 4c), indicating that treatment resulted in increased glucose oxidation in addition to fatty acid oxidation. The rate-limiting enzyme in mammalian nicotinamide adenine dinucleotide (NAD<sup>+</sup>) biosynthesis, nicotinamide phosphoribosyltransferase (NAMPT), has recently been demonstrated to be expressed in a circadian manner leading to a circadian pattern of NAD<sup>+</sup> production and thus regulation of the NAD<sup>+</sup>-dependent deacetylase SIRT1 (refs 21, 22). We found that SR9011 suppressed the circadian rhythm of *Nampt* gene expression

in the liver (Supplementary Fig. 15a), indicating that REV-ERB agonist treatment may alter the post-translational acetylation of proteins that may contribute to some of the physiological alterations that we observe.

In contrast to muscle where there was amplification of the circadian expression of genes coupled to fatty acid oxidation and glycolysis, in white adipose tissue (WAT) we observed a suppression of circadian expression of key genes involved in lipid storage. The expression of both diglyceride acyltransferase 1 and 2 (*Dgat1* and *Dgat2*), the genes encoding the enzyme that catalyses the terminal and committed step in triglyceride synthesis, was suppressed with SR9011 treatment (Fig. 4d). Consistent with this pattern, the circadian expression of another gene involved in triglyceride synthesis, monoacylglycerol acyltransferase (*Mgat1*), is also disturbed (Fig. 4d). Expression of lipid-droplet-associated protein genes including perilipin 1 (*Plin1*) and hormone sensitive lipase (*Hsl*; also called *Lipe*) was also suppressed with SR9011 treatment (Fig. 4d). We also observed suppression of nocturnin (*Ccrn4l*) expression in both adipose and hepatic tissue (Supplementary Fig. 15b). Similar effects on gene expression were noted in mice treated with SR9009 (data not shown).

Clearly, modulation of REV-ERB activity by a synthetic agonist alters the pattern of expression of many genes involved in metabolism in several tissues including the liver, skeletal muscle and WAT. The alterations that we observed are consistent with decreased lipogenesis and cholesterol/bile acid synthesis in the liver, increased lipid and glucose oxidation in the skeletal muscle, and decreased triglyceride synthesis and storage in the WAT.





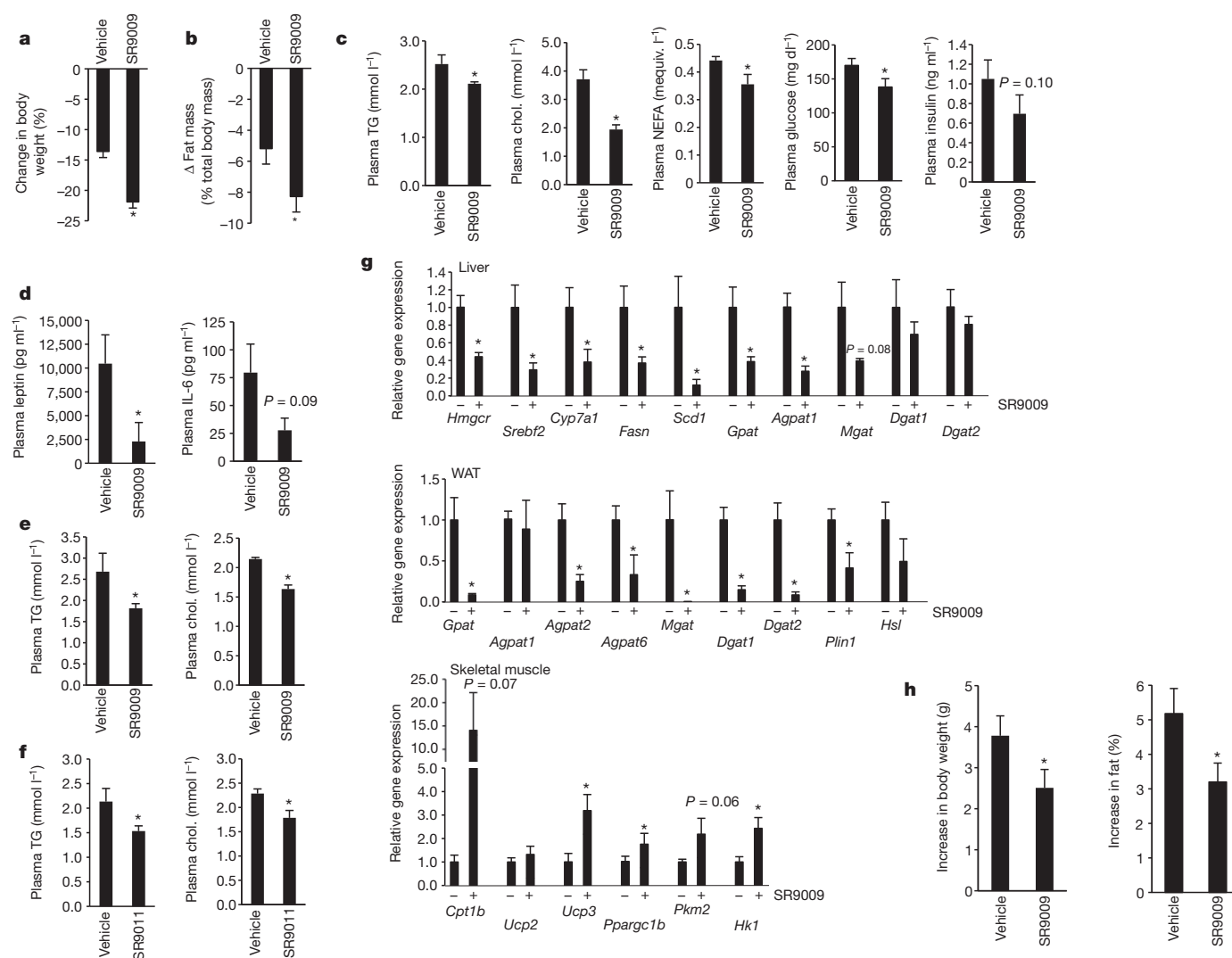
**Figure 4 | REV-ERB ligands alter the pattern of circadian expression of metabolic genes in the liver, skeletal muscle and adipose tissue.** C57BL/6 mice were administered a single dose of SR9011 ( $100 \text{ mg kg}^{-1}$ , i.p.) at CT0 and groups of mice ( $n = 6$ ) were killed and gene expression assessed by QPCR. Graphs were double plotted. **a**, Expression of core clock genes from the liver of vehicle-treated versus SR9011-treated mice. **b**, Expression of metabolic genes

from the liver of vehicle-treated versus SR9011-treated mice. **c**, Expression of metabolic genes from the skeletal muscle of vehicle-treated versus SR9011-treated mice. **d**, Expression of metabolic genes from the white adipose tissue (WAT) of vehicle-treated versus SR9011-treated mice. \* $P < 0.05$ . Error bars indicate mean  $\pm$  s.e.m.;  $n = 6$ –10 mice.

### SR9009 improves the metabolic profile in obese mice

On the basis of the alterations in energy metabolism and gene expression that we observed in normal C57BL/6 and BALB/c mice, we sought to examine whether a REV-ERB- $\alpha/\beta$  agonist would be efficacious in a rodent model of obesity. We initiated the study with 20-week-old C57BL/6 mice (average weight = 41g) that had been maintained on a high-fat diet for 14 weeks (20% carbohydrate 60% fat). The mice continued on the high-fat diet and we initiated twice per day dosing (intraperitoneal (i.p.)) of SR9009. Although the stress of handling and twice-daily injections caused weight loss in vehicle-treated controls, weight loss of SR9009-treated animals was 60% greater (Fig. 5a). During the treatment period, there was no significant difference in the food intake of SR9009- and vehicle-treated mice, although handling itself reduced food intake, explaining the weight loss observed in the controls. SR9009-treated mice exhibited a more severe reduction in adiposity compared with vehicle treatment (Fig. 5b). In addition to the decrease in fat mass we also observed a 12% decrease in plasma triglycerides and a 47% decrease in plasma total cholesterol (Fig. 5c). Plasma non-esterified fatty acids (NEFA) were also reduced (23%) along with plasma glucose (19%) in the SR9009-treated mice (Fig. 5c). There was also a trend towards a decrease in plasma insulin

levels (35%). Consistent with the decrease in adiposity we also noted an 80% decrease in plasma leptin and a decrease (72%) in the proinflammatory cytokine interleukin (IL)-6 (Fig. 5d). Examination of plasma triglycerides and total cholesterol in lean mice also demonstrated the ability of SR9009 and SR9011 to reduce the levels of these lipids (Figs 5e, f). Consistent with the decreased plasma triglyceride levels and total cholesterol levels, we observed a significant decrease in the expression of genes encoding lipogenic enzymes (*Fasn* and *Scd1*) as well as cholesterologenic regulatory proteins (*Hmgcr* and *Srebf2*) with SR9009 treatment (Fig. 5g). In the WAT, SR9009 treatment resulted in a decrease in the expression of genes encoding enzymes involved in triglyceride synthesis (Fig. 5g) as was also observed in lean mice (Fig. 4d). Similar to our observations in lean mice (Fig. 4c), we observed that the REV-ERB agonist induced the expression of genes involved in fatty acid and glucose oxidation (*Cpt1b*, *Ucp3*, *Pparg1b*, *Pkm2* and *Hk1*) (Fig. 5g). Taken together with the results from the CLAMS experiments (Fig. 3), these data indicate that REV-ERB agonists increase energy expenditure by increasing fatty acid and glucose oxidation in the skeletal muscle. The gene expression data are also consistent with decreased triglyceride synthesis in the liver and WAT as well as a reduction in hepatic cholesterol synthesis. We also examined the



**Figure 5 | SR9009 treatment results in a decrease in fat mass and in plasma lipids in diet-induced obese mice.** **a**, Diet-induced obese mice on SR9009 treatment lose weight versus vehicle-treated mice. C57BL/6 mice on a high-fat diet were administered SR9009 (100 mg kg<sup>-1</sup>, i.p., b.i.d., at CT0 and CT12) for 30 days. **b**, Diet-induced obese mice on SR9009 treatment exhibit lower fat mass versus vehicle-treated mice. **c**, Fasting plasma triglycerides (TG), cholesterol (Chol.), non-esterified fatty acids (NEFA) and glucose are decreased in SR9009-treated DIO mice. **d**, Plasma leptin and IL-6 levels from DIO mice. **e**, Fasting plasma triglyceride and cholesterol in lean C57BL/6 mice. Normal mice were

administered 100 mg kg<sup>-1</sup>, i.p., b.i.d. (at CT0 and CT12) SR9009 for 10 days. **f**, Fasting plasma triglyceride and cholesterol are decreased by SR9011 treatment in lean C57BL/6 mice. **g**, Expression of metabolic genes in liver, WAT and skeletal muscle of DIO mice treated with SR9009 as described in **a**. Gene expression was measured by QPCR and normalized to cyclophilin b expression. **h**, SR9009 treatment reduces weight gain in *ob/ob* mice. Body weight and body fat content data are shown from *ob/ob* mice administered SR9009 for 12 days (100 mg kg<sup>-1</sup>, i.p., b.i.d.). \**P* < 0.05. Error bars indicate mean ± s.e.m.; *n* = 6–10 mice.

effects of SR9009 in a genetic model of obesity (*ob/ob* mice) and after 12 days of dosing we observed that SR9009 suppressed the degree of weight gain normally observed in this leptin-deficient mouse with no significant alterations in glucose or insulin tolerance (Fig. 5h and data not shown).

## Discussion

We have developed synthetic REV-ERB-α/β agonists with sufficient pharmacokinetic properties to examine their activity *in vivo*. These compounds alter the circadian pattern of expression of core clock genes as well as circadian locomotor behaviour in mice. A single injection of either SR9011 or SR9009 results in loss of the subsequent active period in mice maintained under constant dark conditions. Consistent with this major alteration in circadian behaviour the circadian pattern of core clock gene expression in the hypothalamus is perturbed. In *Rev-erb-α* null mice circadian locomotor activity is also disturbed, with the mice displaying a decreased period relative to wild-type mice under constant dark or constant light conditions<sup>4</sup>.

We would not necessarily expect to mimic the phenotype of either constant over- or under-expression of REV-ERB with a pharmacological REV-ERB ligand, as modulation of the receptors' activity would be only transient. This is also probably the reason that on the day after administration of the REV-ERB agonist, normal circadian behaviour is completely restored. We also administered the REV-ERB agonists under L/D conditions to mimic a therapeutic situation as well as the metabolic studies where mice were maintained on this standard L/D cycle. In this case the effects of administration of the REV-ERB ligands were considerably less severe both in terms of alterations in patterns of core clock gene expression in the hypothalamus and in circadian locomotor behaviour. A single injection of either SR9009 or SR9011 resulted in a 1–3-h delay in initiation of diurnal activity. Consistent with this observation, when we examined oxygen consumption we observed a similar delay in the nocturnal peak in VO<sub>2</sub>. Thus, synthetic REV-ERB ligands effectively alter the physiological time of day of mice, indicating that this class of compound may be useful for the treatment of sleep disorders. Additionally, synthetic



REV-ERB ligands may hold utility for jet lag where the compounds could be used to realign both the central and peripheral clocks to a new time zone.

The core clock machinery is closely associated with metabolic regulation and there are a myriad of examples of genetic alterations to clock genes leading to metabolic disturbances and even metabolic diseases in rodent models<sup>2,18,23–25</sup>. In addition to its role in direct modulation of the positive arm of the mammalian circadian oscillator, REV-ERB has also been demonstrated to have a direct role in the regulation of an array of metabolic genes<sup>26–28</sup>. Our observations that the REV-ERB agonists increase energy expenditure and decrease fat mass, plasma triglycerides and cholesterol levels, suggest that these compounds may be useful in the treatment of metabolic diseases.

## METHODS SUMMARY

**Synthesis of SR9009 and SR9011.** Compounds were synthesized by reductive amination of 5-nitro-2-thiophenecarboxaldehyde with 4-chlorobenzylamine, and sodium triacetoxyborohydride yielded the secondary amine. A second reductive amination with 1-Boc-pyrrolidine-3-carboxaldehyde yielded the tertiary amine. This compound was treated with trifluoroacetic acid to remove the Boc-protecting group, and then reacted with either ethyl chloroformate (SR9009) or pentyl isocyanate (SR9011) to give the desired products.

**Cell culture and co-transfections.** HEK293 cells were maintained and transfected with vectors previously described<sup>12</sup>.

**Mouse studies.** For circadian gene expression experiments, male C57BL/6 mice (8–10 weeks of age) were either maintained on a L/D (12 h/12 h) cycle or on constant darkness. At circadian time (CT) 0 mice were administered a single dose of 100 mg kg<sup>−1</sup> SR9009 or SR9011 (i.p.) and groups of animals (*n* = 6) were killed at CT0, CT6, CT12 and CT18. Gene expression was determined by real-time QPCR.

Received 14 June 2011; accepted 12 March 2012.

Published online 29 March 2012.

1. Welsh, D. K., Takahashi, J. S. & Kay, S. A. Suprachiasmatic nucleus: Cell autonomy and network properties. *Annu. Rev. Physiol.* **72**, 551–577 (2010).
2. Bass, J. & Takahashi, J. S. Circadian integration of metabolism and energetics. *Science* **330**, 1349–1354 (2010).
3. Ko, C. H. & Takahashi, J. S. Molecular components of the mammalian circadian clock. *Hum. Mol. Genet.* **15**, R271–R277 (2006).
4. Preitner, N. *et al.* The orphan nuclear receptor rev-erb $\alpha$  controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251–260 (2002).
5. Crumbley, C. & Burris, T. P. Direct regulation of CLOCK expression by REV-ERB. *PLoS ONE* **6**, e17290 (2011).
6. Raghuram, S. *et al.* Identification of heme as the ligand for the orphan nuclear receptors REV-ERB $\alpha$  and REV-ERB $\beta$ . *Nature Struct. Mol. Biol.* **14**, 1207–1213 (2007).
7. Yin, L. *et al.* Rev-erb $\alpha$ , a heme sensor that coordinates metabolic and circadian pathways. *Science* **318**, 1786–1789 (2007).
8. Kojetin, D., Wang, Y., Kamenecka, T. M. & Burris, T. P. Identification of sr8278, a synthetic antagonist of the nuclear heme receptor REV-ERB. *ACS Chem. Biol.* **6**, 131–134 (2011).
9. Kumar, N. *et al.* Regulation of adipogenesis by natural and synthetic REV-ERB ligands. *Endocrinology* **151**, 3015–3025 (2010).
10. Grant, D. *et al.* Gsk4112, a small molecule chemical probe for the cell biology of the nuclear heme receptor Rev-erb $\alpha$ . *ACS Chem. Biol.* **5**, 925–932 (2010).
11. Meng, Q. J. *et al.* Ligand modulation of REV-ERB $\alpha$  function resets the peripheral circadian clock in a phasic manner. *J. Cell Sci.* **121**, 3629–3635 (2008).
12. Kumar, N. *et al.* The benzenesulfoamide t0901317 [N-(2,2,2-trifluoroethyl)-N-[4-[2,2,2-trifluoro-1-hydroxy-1-(trifluoromethyl)ethyl]phenyl]-benzenesulfonamide] is a novel retinoic acid receptor-related orphan receptor- $\alpha/\gamma$  inverse agonist. *Mol. Pharmacol.* **77**, 228–236 (2010).
13. Busby, S. A. *et al.* Identification of a novel non-retinoid pan inverse agonist of the retinoic acid receptors. *ACS Chem. Biol.* **6**, 618–627 (2011).
14. Yoo, S. H. *et al.* Period2: Luciferase real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc. Natl Acad. Sci. USA* **101**, 5339–5346 (2004).
15. Wang, J., Yin, L. & Lazar, M. A. The orphan nuclear receptor Rev-erb $\alpha$  regulates circadian expression of plasminogen activator inhibitor type 1. *J. Biol. Chem.* **281**, 33842–33848 (2006).
16. Le Martelat, G. *et al.* REV-ERB $\alpha$  participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biol.* **7**, e1000181 (2009).
17. Duez, H. *et al.* Regulation of bile acid synthesis by the nuclear receptor Rev-erb $\alpha$ . *Gastroenterology* **135**, 689–698 (2008).
18. Green, C. B., Takahashi, J. S. & Bass, J. The meter of metabolism. *Cell* **134**, 728–742 (2008).
19. Ramakrishnan, S. N., Lau, P., Burke, L. J. & Muscat, G. E. O. Rev-erb $\beta$  regulates the expression of genes involved in lipid absorption in skeletal muscle cells—evidence for cross-talk between orphan nuclear receptors and myokines. *J. Biol. Chem.* **280**, 8651–8659 (2005).
20. Raspe, E. *et al.* Identification of Rev-erb $\alpha$  as a physiological repressor of apoC-III gene transcription. *J. Lipid Res.* **43**, 2172–2179 (2002).
21. Ramsey, K. M. *et al.* Circadian clock feedback cycle through NAMPT-mediated NAD<sup>+</sup> biosynthesis. *Science* **324**, 651–654 (2009).
22. Nakahata, Y., Sahar, S., Astarita, G., Kaluzova, M. & Sassone-Corsi, P. Circadian control of the NAD<sup>+</sup> salvage pathway by CLOCK-SIRT1. *Science* **324**, 654–657 (2009).
23. Bray, M. S. & Young, M. E. Regulation of fatty acid metabolism by cell autonomous circadian clocks: time to fatten up on information? *J. Biol. Chem.* **286**, 11883–11889 (2011).
24. Asher, G. & Schibler, U. Crosstalk between components of circadian and metabolic cycles in mammals. *Cell Metab.* **13**, 125–137 (2011).
25. Gimble, J. M., Sutton, G. M., Bunnell, B. A., Ptitsyn, A. A. & Floyd, Z. E. Prospective influences of circadian clocks in adipose tissue and metabolism. *Nature Rev. Endocrinol.* **7**, 98–107 (2011).
26. Duez, H. & Staels, B. Rev-erb- $\alpha$ : An integrator of circadian rhythms and metabolism. *J. Appl. Physiol.* **107**, 1972–1980 (2009).
27. Burris, T. P. Nuclear hormone receptors for heme: REV-ERB $\alpha$  and REV-ERB $\beta$  are ligand-regulated components of the mammalian clock. *Mol. Endocrinol.* **22**, 1509–1520 (2008).
28. Yin, L., Wu, N. & Lazar, M. A. Nuclear receptor Rev-erb $\alpha$ : a heme receptor that coordinates circadian rhythm and metabolism. *Nucl. Recept. Signal.* **8**, e001 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by National Institutes of Health Grants (DK080201, MH092769 and DK089984) and the Howard Hughes Medical Institute. L.A.S. is the recipient of an individual National Research Service Award (DK088499).

**Author Contributions** T.P.B. conceived the project. R.N., Y.S. and T.M.K. synthesized and analysed the ligands. L.A.S., Y.W., S.B., D.J.K. and T.P.B. designed/analysed and/or performed the transfection and biochemical assays. D.J.K. and T.H. designed and performed the CD analysis. T.P.B., L.A.S. and A.A.B. designed, analysed and performed the metabolic studies. Y.W., T.P.B., S.B. and T.L. designed/analysed and/or performed the circadian gene expression and behaviour analysis. J.L. performed gene expression analysis. S.-H.Y. and J.S.T. designed and performed the studies using the *Per2-luc* mouse tissues. M.D.C. performed the pharmacokinetic analysis. T.P.B. wrote the manuscript with input from all of the authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to T.P.B. (tburris@scripps.edu).

# Osteoprotection by semaphorin 3A

Mikihito Hayashi<sup>1,2,3</sup>, Tomoki Nakashima<sup>1,2,3</sup>, Masahiko Taniguchi<sup>4</sup>, Tatsuhiko Kodama<sup>5</sup>, Atsushi Kumanogoh<sup>6,7</sup> & Hiroshi Takayanagi<sup>1,2,3,8</sup>

**The bony skeleton is maintained by local factors that regulate bone-forming osteoblasts and bone-resorbing osteoclasts, in addition to hormonal activity. Osteoprotegerin protects bone by inhibiting osteoclastic bone resorption, but no factor has yet been identified as a local determinant of bone mass that regulates both osteoclasts and osteoblasts. Here we show that semaphorin 3A (Sema3A) exerts an osteoprotective effect by both suppressing osteoclastic bone resorption and increasing osteoblastic bone formation. The binding of Sema3A to neuropilin-1 (Nrp1) inhibited receptor activator of nuclear factor- $\kappa$ B ligand (RANKL)-induced osteoclast differentiation by inhibiting the immunoreceptor tyrosine-based activation motif (ITAM) and RhoA signalling pathways. In addition, Sema3A and Nrp1 binding stimulated osteoblast and inhibited adipocyte differentiation through the canonical Wnt/ $\beta$ -catenin signalling pathway. The osteopenic phenotype in *Sema3a*<sup>-/-</sup> mice was recapitulated by mice in which the Sema3A-binding site of Nrp1 had been genetically disrupted. Intravenous Sema3A administration in mice increased bone volume and expedited bone regeneration. Thus, Sema3A is a promising new therapeutic agent in bone and joint diseases.**

Bone homeostasis has long been thought to be predominantly maintained by the endocrine system by calcium regulating hormones, but increasing evidence indicates that bone is also under the control of factors related to immune and neuronal regulation<sup>1,2</sup>. As an imbalance between bone resorption and formation results in metabolic bone disorders such as osteoporosis<sup>3,4</sup>, understanding the balancing mechanisms is important for the development of therapeutic agents. Because bone formation is linked to resorption through coupling factors<sup>5,6</sup>, treatment with anti-resorptive agents results in simultaneous suppression of bone formation, with the result that the efficacy is compromised<sup>6-9</sup>. It is thus crucial to identify a molecule that can regulate both resorption and formation synchronously.

Osteoclasts are derived from monocyte/macrophage precursor cells, and their differentiation is regulated by mesenchymal cells, such as osteoblasts, chondrocytes and osteocytes, which express the key osteoclast differentiation factor RANKL<sup>1,10-12</sup>. Osteoblastic cells counterbalance the function of RANKL by producing a soluble decoy receptor for RANKL, osteoprotegerin (Opg), the name of which indicates it is a protector of bone<sup>13</sup>. Here we show that conditioned medium from Opg-deficient mouse calvarial cells contains factors that inhibit osteoclast formation, and one of these factors is the axon guidance molecule Sema3A. Sema3A inhibits osteoclast differentiation and promotes osteoblastic bone formation, and is thus a potent osteoprotective factor produced by osteoblastic cells.

## Sema3A mediates anti-osteoclastogenesis in osteoblasts

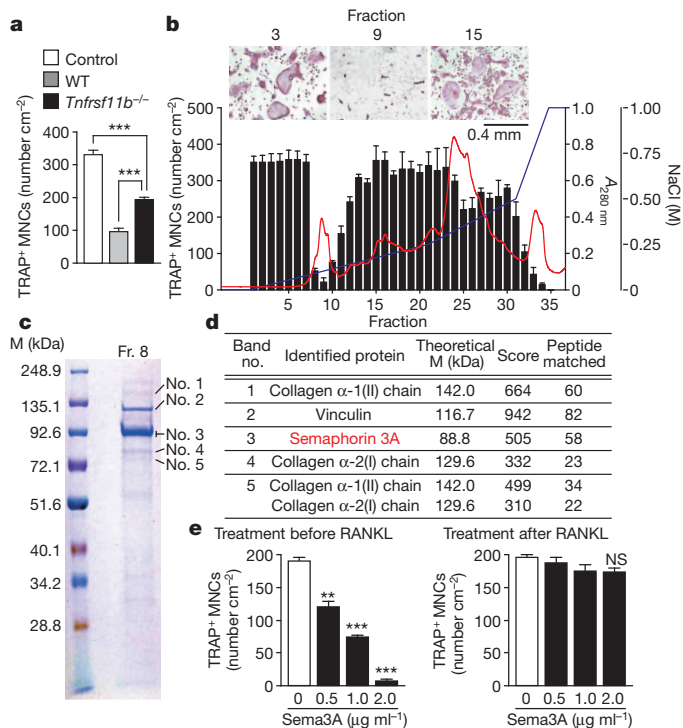
The conditioned medium of osteoblastic cells was able to inhibit osteoclast differentiation of bone marrow-derived monocyte/macrophage precursor cells (BMMs) stimulated by RANKL in the presence of macrophage colony-stimulating factor (M-CSF) (Fig. 1a). We observed a substantial anti-osteoclastogenic effect in the conditioned medium of calvarial cells lacking the *Tnfrsf11b* gene (encoding Opg in mice) (Fig. 1a), suggesting the presence of one or more other soluble

inhibitory factors. To identify the osteoblast-secreted proteins that inhibit osteoclast differentiation, we fractionated the conditioned medium of *Tnfrsf11b*<sup>-/-</sup> calvarial cells by anion-exchange liquid chromatography (Fig. 1b). We found that fractions 8–10 exerted a potent inhibitory effect on osteoclast differentiation (Fig. 1b). The proteins in fraction 8 were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Fig. 1c), and the major bands were excised and analysed by liquid chromatography–tandem mass spectrometry (LC-MS/MS). Among the identified proteins (Fig. 1d), we focused on the axon guidance molecule Sema3A<sup>14,15</sup> (Supplementary Fig. 1a), as recent studies have suggested that axon guidance molecules are involved in the interaction between osteoblasts and osteoclasts<sup>16-18</sup>. Previous reports suggest that Sema3A expressed in the skeletal system may have a role in the regulation of innervation and blood vessel invasion and contribute to skeletal patterning<sup>19-21</sup>, but the function of Sema3A in the regulation of bone remodelling remains unknown.

Western blot analysis confirmed that Sema3A protein was present in fractions 8 and 9 (Supplementary Fig. 1b), and that the inhibitory effect of fraction 9 is largely mediated by Sema3A, as this effect was abrogated by the addition of soluble Nrp1, which functioned as a Sema3A decoy receptor (Supplementary Fig. 1c). The addition of recombinant Sema3A potentially inhibited osteoclast differentiation in a dose-dependent manner when Sema3A was added before RANKL treatment (Fig. 1e). When Sema3A was added after RANKL treatment, the inhibitory effect was not observed (Fig. 1e). *Sema3a* was predominantly expressed in osteoblast lineage cells among various cells examined (Supplementary Fig. 2a), whereas Sema3A was not detected in osteoclasts (Supplementary Fig. 2a–c). The expression of *Sema3a* messenger RNA in isolated osteocytes and osteoblasts was comparable (Supplementary Fig. 2d). *Sema3a* mRNA expression in calvarial cells was higher than that of any other semaphorin family member tested (Supplementary Fig. 2e).

<sup>1</sup>Department of Cell Signaling, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Yushima 1-5-45, Bunkyo-ku, Tokyo 113-8549, Japan. <sup>2</sup>Japan Science and Technology Agency, Exploratory Research for Advanced Technology Program, Takayanagi Osteonetwork Project, Yushima 1-5-45, Bunkyo-ku, Tokyo 113-8549, Japan. <sup>3</sup>Global Center of Excellence Program, International Research Center for Molecular Science in Tooth and Bone Diseases, Yushima 1-5-45, Bunkyo-ku, Tokyo 113-8549, Japan. <sup>4</sup>Department of Molecular Medical Sciences, Research Institute for Frontier Medicine, Sapporo Medical University School of Medicine, S-1, W-17, Chuo-ku, Sapporo 060-8556, Japan. <sup>5</sup>Laboratory for Systems Biology and Medicine, Research Center for Advanced Science and Technology, Department of Molecular Biology and Medicine, University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153-8904, Japan. <sup>6</sup>Department of Respiratory Medicine, Allergy and Rheumatic Diseases, Graduate School of Medicine, Osaka University, Yamadaoka 2-2, Suita, Osaka 565-0871, Japan. <sup>7</sup>Department of Immunopathology, Immunology Frontier Research Center, Osaka University, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan. <sup>8</sup>Centre for Orthopaedic Research, School of Surgery, The University of Western Australia, Nedlands, Western Australia 6009, Australia.





**Figure 1 | Identification of Semaphorin 3A as an inhibitory factor of osteoclast differentiation.** **a**, Effect of wild-type (WT) and Opg-deficient (*Tnfrsf11b*<sup>-/-</sup>) calvarial cell-conditioned medium on osteoclast differentiation. MNC, mononuclear cells. **b**, Fractionation of Opg-deficient calvarial cell-conditioned medium by anion-exchange chromatography, and the effect of each fraction on osteoclast differentiation (black bars). Absorbance at 280 nm is indicated as a red line and the concentration of NaCl as a blue line. Inset shows TRAP staining of osteoclast cultures treated with fractions 3, 9 and 15. **c**, Coomassie brilliant blue-stained SDS-PAGE image of fraction 8. M, molecular mass. **d**, List of the identified proteins (that is, more than 300 on the MASCOT score), their theoretical molecular mass, MASCOT score and the number of non-redundant peptides. **e**, Effect of Semaphorin 3A treatment on osteoclast differentiation. Error bars (**a**, **b** and **e**) denote mean  $\pm$  s.e.m. \*\**P* < 0.01; \*\*\**P* < 0.005; NS, not significant.

## Semaphorin 3A regulates osteoclast differentiation via Nrp1

Microcomputed tomography along with bone morphometric and radiographic analyses showed that *Sema3a*<sup>-/-</sup> mice<sup>19</sup> have a severe osteopenic phenotype both in trabecular and cortical bones, accompanied by an increase in the osteoclast number and eroded surface (Fig. 2a, b and Supplementary Fig. 3a–c). There was no difference in the number of osteoclast precursor CD11b<sup>low/-</sup>CD3ε<sup>-</sup>B220<sup>-</sup>c-fms<sup>+</sup>c-kit<sup>+</sup> cells<sup>22</sup> in the bone marrow of wild-type and *Sema3a*<sup>-/-</sup> mice (Supplementary Fig. 3d).

When osteoclast formation was analysed in a coculture of bone marrow and calvarial cells, the formation of tartrate-resistant acid phosphatase (TRAP)-positive multinucleated osteoclasts was markedly enhanced in *Sema3a*<sup>-/-</sup> cells (Fig. 2c and Supplementary Fig. 3e, f). This enhanced osteoclastogenesis was not observed when BMMs were stimulated by RANKL and M-CSF (Supplementary Fig. 3g). In addition, enhanced osteoclastogenesis was observed in the coculture of wild-type bone marrow cells and *Sema3a*<sup>-/-</sup> calvarial cells, but not in the coculture of *Sema3a*<sup>-/-</sup> bone marrow cells and wild-type calvarial cells (Fig. 2c and Supplementary Fig. 3e, f). These results indicate that the osteoblastic expression of Semaphorin 3A inhibits osteoclastogenesis. The level of RANKL and Opg in the calvarial cells or the serum was not affected by Semaphorin 3A deficiency (Supplementary Fig. 3h, i).

Semaphorin 3A binds to a receptor complex of the ligand-binding subunit Nrp1 and one of the class A plexins (PlxnA1, PlxnA2, PlxnA3 and PlxnA4), which function as the signal-transducing subunit<sup>23</sup>. We found that Nrp1 expression in BMMs was rapidly and markedly

suppressed after RANKL stimulation (Fig. 2d). Because Semaphorin 3A-induced inhibition of osteoclastogenesis was observed only when Semaphorin 3A was added before RANKL stimulation (Fig. 1e), we proposed that Semaphorin 3A does not inhibit osteoclastogenesis after RANKL stimulation owing to Nrp1 downregulation.

When Nrp1 was overexpressed by retroviral transfer, Semaphorin 3A exerted an inhibitory effect even when Semaphorin 3A had been added after RANKL stimulation (Supplementary Fig. 3j). Notably, osteoclastogenesis was inhibited by Nrp1 overexpression only (Supplementary Fig. 3j). When *Nrp1* expression was knocked down by short hairpin RNA (shRNA), the inhibitory effect of Semaphorin 3A on osteoclast differentiation was abolished (Supplementary Fig. 3k). Thus, the level of Nrp1 correlates with the inhibitory effect of Semaphorin 3A on osteoclastogenesis, suggesting that the Nrp1 downregulation caused by RANKL signalling is important for proper osteoclast differentiation by cancelling the inhibitory effect of Semaphorin 3A.

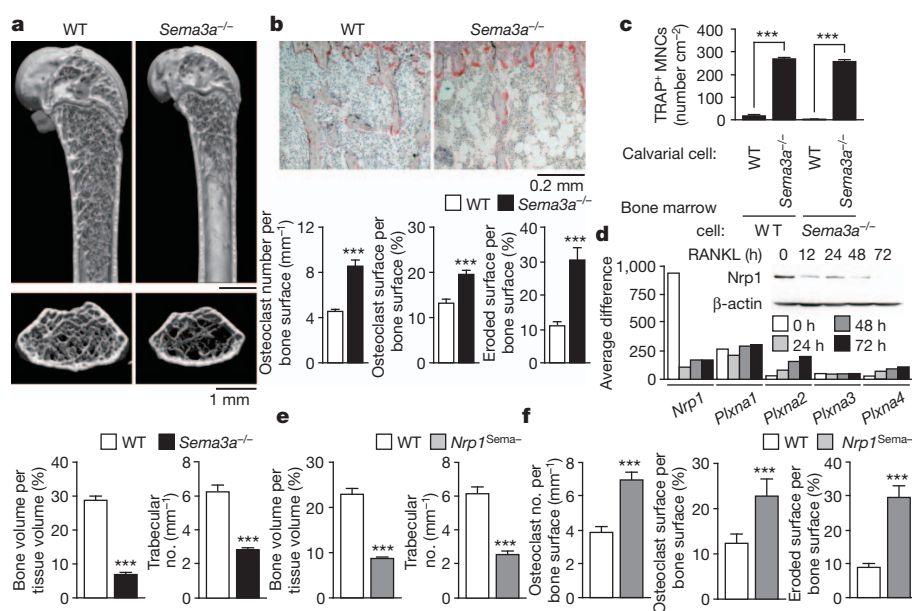
Furthermore, we analysed knockin mice in which the *Nrp1* gene was replaced by mutant *Nrp1* lacking the Sema-binding site (*Nrp1*<sup>Sema-</sup> mice), as *Nrp1*<sup>-/-</sup> mice are embryonically lethal and Nrp1 also contains the vascular endothelial growth factor (VEGF)-binding site<sup>24</sup>. Semaphorin 3A did not inhibit RANKL-induced osteoclast differentiation in *Nrp1*<sup>Sema-</sup> cells (Supplementary Fig. 3l), showing that Semaphorin 3A inhibits osteoclastogenesis by binding to Nrp1. As expected, *Nrp1*<sup>Sema-</sup> mice showed an osteopenic phenotype accompanied by enhanced osteoclast differentiation, which was similar to *Sema3a*<sup>-/-</sup> mice (Fig. 2e, f and Supplementary Fig. 3m–o).

To understand the mechanism of the RANKL-induced inhibition of Nrp1 expression, we examined the involvement of the transcription factors nuclear factor- $\kappa$ B (NF- $\kappa$ B), c-Fos and nuclear factor of activated T cells c1 (NFATc1), which are all activated by RANKL<sup>1</sup>. RANKL-induced downregulation of *Nrp1* expression was abolished by an NF- $\kappa$ B inhibitor, but was not affected by the deficiency of NFATc1 or c-Fos (Supplementary Fig. 4a–c). Chromatin immunoprecipitation analysis showed that NF- $\kappa$ B p65 and, to a lesser extent, p50 were recruited to the proximal NF- $\kappa$ B-binding site of the *Nrp1* promoter after RANKL stimulation (Supplementary Fig. 4d, e). NF- $\kappa$ B p65 and p50 inhibited *Nrp1* promoter activity in a reporter gene assay (Supplementary Fig. 4f). Retroviral overexpression of p65 in BMMs led to Nrp1 downregulation in the absence of RANKL, which was further facilitated by p50 overexpression (Supplementary Fig. 4g). These effects were dependent on histone deacetylases (Supplementary Fig. 4g, h), suggesting that the recruitment of corepressors<sup>25</sup> by RANKL-stimulated NF- $\kappa$ B is involved in Nrp1 downregulation.

## Mechanism of anti-osteoclastogenesis by Semaphorin 3A

The binding of RANKL to its receptor RANK results in the activation of tumour-necrosis factor (TNF) receptor-associated factor 6 (TRAF6), which stimulates the NF- $\kappa$ B and mitogen-activated protein kinase (MAPK) pathways<sup>1</sup>. RANKL also activates the activator protein 1 (AP-1) transcription factor complex, including c-Fos, which cooperates with NF- $\kappa$ B to induce NFATc1, thus activating the transcription of osteoclast-specific genes<sup>1,26</sup>. The robust induction of NFATc1 is dependent on calcium signalling (costimulatory signalling) stimulated by the ITAM-bearing adaptor molecules DNAX-activation protein 12 (DAP12) and Fc receptor common  $\gamma$  subunit (FcR $\gamma$ ), which associate with immunoglobulin-like receptors such as triggering receptor expressed on myeloid cells 2 (TREM2) and osteoclast-associated receptor (OSCAR)<sup>1,27</sup>.

From these observations, the question arises as to how the Semaphorin 3A–Nrp1 axis inhibits osteoclastogenic signalling. RANKL-stimulated induction of the osteoclastic genes *Ctsk*, *Acp5* and *Nfatc1* was severely impaired by Semaphorin 3A without affecting the expression of *Tnfrsf11a* or *Csf1r* (Fig. 3a). There was no difference in the cell proliferation rate or the percentage of apoptotic cells among the osteoclast precursor cells between the Semaphorin 3A-treated and the control cells (Supplementary Fig. 5a, b). RANKL-induced activation



**Figure 2** | *Sema3a*<sup>-/-</sup> and *Nrpl1*<sup>Sema-</sup> mice show a severe low bone mass phenotype.

**a**, Microcomputed tomography images of the femurs of 10-week-old *Sema3a*<sup>-/-</sup> mice and their wild-type littermates ( $n = 4-6$ ). The bone volume and parameters of trabecular bone were determined by microcomputed tomography analysis. **b**, TRAP staining of the proximal tibiae of *Sema3a*<sup>-/-</sup> mice and their wild-type littermates ( $n = 4-6$ ). Osteoclastic parameters were measured using bone morphometric analysis. **c**, Osteoclast differentiation from wild-type or *Sema3a*<sup>-/-</sup> bone marrow cells in coculture with wild-type or *Sema3a*<sup>-/-</sup> calvarial cells at day 4. **d**, GeneChip analysis of the mRNA expression of *Nrpl1* and *Plxn1-4* during osteoclast differentiation. *Nrpl1* protein expression in BMMs stimulated with RANKL was analysed by western blot (inset). **e**, Microcomputed tomography analysis of the femurs of 10-week-old *Nrpl1*<sup>Sema-</sup> mice and their wild-type littermates ( $n = 4-5$ ). **f**, Parameters for osteoclastic bone resorption in the bone morphometric analysis of the proximal tibiae of *Nrpl1*<sup>Sema-</sup> mice and their wild-type littermates ( $n = 4-5$ ). Error bars (**a-c**, **e** and **f**) denote mean  $\pm$  s.e.m. \*\*\* $P < 0.01$ ; \*\*\*\* $P < 0.005$ .

of the signalling pathways downstream of TRAF6, including the MAPKs (such as ERK, JNK, p38) and inhibitor of  $\kappa$ B (I $\kappa$ B) kinases, was comparable in BMMs with or without Sema3A (Supplementary Fig. 5c).

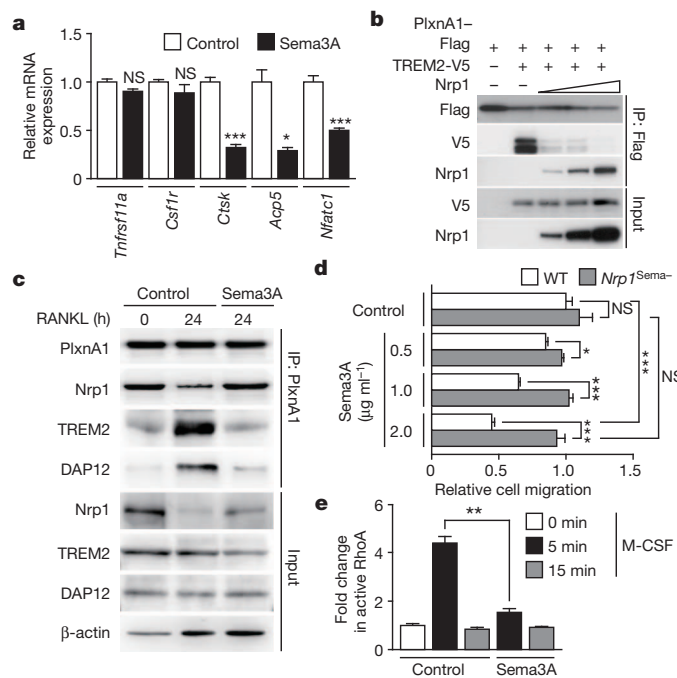
PlxnA1 promotes osteoclast differentiation by activating the ITAM signal through the formation of the PlxnA1-TREM2-DAP12 complex in response to ligands such as Sema6D<sup>17</sup>. However, PlxnA1 is constitutively associated with *Nrpl1*, which mediates Sema3A signalling

instead of TREM2-DAP12 signalling<sup>28</sup>. With increasing *Nrpl1* expression, the amount of TREM2 associated with PlxnA1 decreased and *Nrpl1* associated with PlxnA1 increased (Fig. 3b).

RANKL induced the formation of the PlxnA1-TREM2-DAP12 complex by the downregulation of *Nrpl1*, thereby releasing PlxnA1 from the PlxnA1-Nrpl1 complex (Fig. 3c). Sema3A treatment inhibited RANKL-induced formation of the PlxnA1-TREM2-DAP12 complex by inhibiting *Nrpl1* downregulation and maintaining the PlxnA1-Nrpl1 complex (Fig. 3c). The cell surface and intracellular expression of *Nrpl1* was highly downregulated by RANKL treatment (Supplementary Fig. 5d), but Sema3A treatment induced the internalization of *Nrpl1*, as already reported<sup>29</sup>, and protected RANKL-induced *Nrpl1* downregulation without altering *Nrpl1* mRNA expression (Supplementary Fig. 5d, e).

RANKL-induced tyrosine phosphorylation of phospholipase C $\gamma$ 2 (PLC $\gamma$ 2) and calcium oscillation were both markedly blocked by Sema3A treatment (Supplementary Fig. 5f, g). We observed that osteoclast differentiation in DAP12-deficient (also known as *Tyrbp*-deficient) bone marrow cells was not enhanced even in a coculture with *Sema3a*<sup>-/-</sup> calvarial cells (Supplementary Fig. 5h), suggesting that Sema3A-induced inhibition is mediated by the modulation of DAP12-induced ITAM signalling. Thus, *Nrpl1* competes with TREM2 for PlxnA1, thereby functioning as a suppressor of the PlxnA1-TREM2-DAP12-induced costimulatory signal. Sema3A-induced inhibition of osteoclast differentiation was less observed in the presence of Sema6D (Supplementary Fig. 5i).

We examined the effect of Sema3A on the migration of BMMs because the semaphorin-plexin system regulates actin cytoskeletal rearrangement<sup>15,23</sup>. We observed a repulsive effect of Sema3A on M-CSF-induced migration of BMMs (Fig. 3d). By contrast, this repulsive effect was not observed in *Nrpl1*<sup>Sema-</sup> BMMs (Fig. 3d). Because semaphorin-plexin signalling regulates the Rho family of small GTPases<sup>15,23</sup>, we examined the effect of Sema3A treatment on M-CSF-induced activation of the RhoA and Rac GTPases. Sema3A treatment abrogated RhoA activation in response to M-CSF (Fig. 3e), but not Rac activation (Supplementary Fig. 5j), suggesting that the inhibition of RhoA activation is involved in the inhibitory effect of Sema3A on the migration of BMMs.



**Figure 3** | Inhibition of osteoclast differentiation by Sema3A-Nrpl1 signalling. **a**, Effect of Sema3A treatment on osteoclastic gene expression in BMMs treated with RANKL for 2 days. **b**, Effect of *Nrpl1* expression on the association of PlxnA1 with TREM2. IP, immunoprecipitation. **c**, Effect of Sema3A treatment on the formation of a complex of PlxnA1 with *Nrpl1* or TREM2/DAP12 in RANKL-treated BMMs. **d**, Transwell assay of the effect of Sema3A on M-CSF-induced migration of BMMs derived from wild-type or *Nrpl1*<sup>Sema-</sup> mice. **e**, Effect of Sema3A treatment on the activation of RhoA in BMMs stimulated with M-CSF. Error bars (**a**, **d** and **e**) denote mean  $\pm$  s.e.m. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.005$ .

## Sema3A regulates osteoblasts through the Wnt pathway

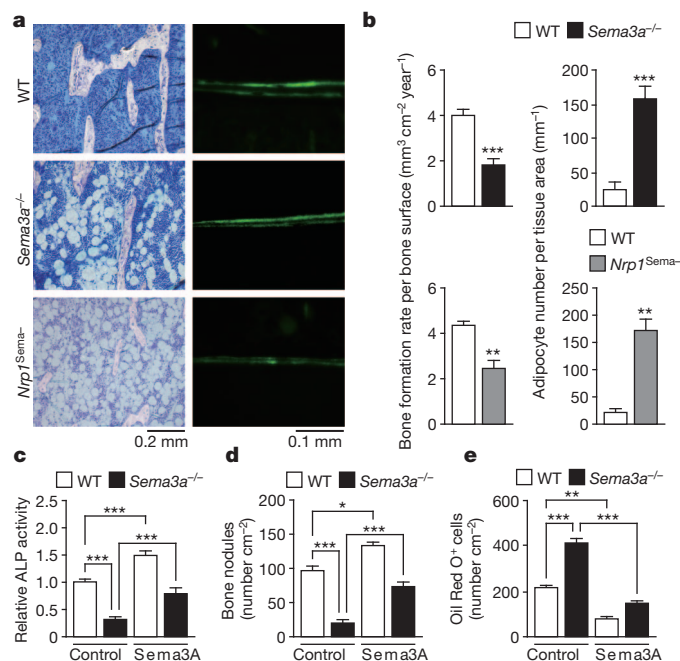
In addition to an osteoclastic phenotype, both *Sema3a*<sup>-/-</sup> and *Nrpl1*<sup>Sema-</sup> mice were found to have osteoblastic and adipocytic



phenotypes (Fig. 4a and Supplementary Fig. 6a–c); that is, they had a decreased osteoblast number, a reduced bone formation rate and a markedly increased adipocyte number (Fig. 4b and Supplementary Fig. 6d, e) without any significant difference in the weight of the epididymal white adipose tissue per body weight (Supplementary Fig. 6f, g). Taken together, the severe osteopenic phenotype in *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> mice was caused by both a decrease in the osteoblastic bone formation and an increase in osteoclastic bone resorption.

Calvarial cells obtained from *Sema3a*<sup>-/-</sup> or *Nrpl*<sup>Sema</sup> mice were cultured in an osteogenic medium with or without Sema3A. Alkaline phosphatase activity and bone nodule formation were markedly decreased in *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> cells, and Sema3A treatment facilitated the differentiation of *Sema3a*<sup>-/-</sup> calvarial cells into osteoblastic cells, but not *Nrpl*<sup>Sema</sup> cells (Fig. 4c, d and Supplementary Fig. 6h–k). Neither the cell proliferation rate nor the percentage of apoptotic cells was affected in both types of mutant cells (Supplementary Fig. 6l, m). Adipocyte differentiation was highly increased in both *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> cells, and Sema3A treatment blocked the differentiation of wild-type and *Sema3a*<sup>-/-</sup> cells into adipocytes, but not *Nrpl*<sup>Sema</sup> cells (Fig. 4e and Supplementary Fig. 6n, o). In *Sema3a*<sup>-/-</sup> cells, the expression of the osteoblast genes *Runx2*, *Sp7* (which encodes osteix), *Alpl* and *Bglap* (encoding osteocalcin) was strongly suppressed (Supplementary Fig. 7a), and the expression of the adipocyte genes *Pparg*, *Cebpa*, *Fabp4* (encoding aP2) and *Lpl* (encoding lipoprotein lipase) was highly increased (Supplementary Fig. 7b). These results indicate that Sema3A activates osteoblast differentiation and inhibits adipocyte differentiation through Nrpl.

Because the mRNA expression levels of the known regulators of mesenchymal cell differentiation<sup>30,31</sup> were comparable in wild-type and *Sema3a*<sup>-/-</sup> cells (Supplementary Fig. 7c), we performed gene



**Figure 4 | Impaired osteoblast differentiation and increased adipocyte differentiation in *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> mice.** **a**, Toluidine blue staining of the proximal tibiae of wild-type, *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> mice (left). New bone formation was determined by calcein double labelling (right). **b**, Osteoblastic and adipocytic parameters measured by histomorphometric analysis of wild-type, *Sema3a*<sup>-/-</sup> and *Nrpl*<sup>Sema</sup> mice ( $n = 4-6$ ). **c**, Alkaline phosphatase (ALP) staining of wild-type and *Sema3a*<sup>-/-</sup> calvarial cells cultured in osteogenic medium with or without Sema3A. **d**, Bone nodule formation in wild-type and *Sema3a*<sup>-/-</sup> calvarial cells cultured in osteogenic medium with or without Sema3A. **e**, Adipocyte differentiation in wild-type and *Sema3a*<sup>-/-</sup> bone marrow stromal cells cultured in adipogenic medium with or without Sema3A. Error bars (b–e) denote mean  $\pm$  s.e.m. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.005$ .

expression profiling of the calvarial cells derived from *Sema3a*<sup>-/-</sup> mice to obtain insight into the Sema3A-activated molecular pathways in osteoblasts. Gene set enrichment analysis in *Sema3a*<sup>-/-</sup> cells showed a significant downregulation of the gene sets involved in the Wnt signalling pathway and the Wnt-related signalling pathways (Supplementary Fig. 7d and Supplementary Table 1). We therefore focused on the canonical Wnt pathway, as it is known to promote osteoblast differentiation and inhibit adipocyte differentiation<sup>31–33</sup>. The mRNA expression of most of the transcriptional targets of  $\beta$ -catenin was considerably reduced (Supplementary Fig. 7e), and the Wnt3a-induced nuclear accumulation of  $\beta$ -catenin was suppressed in *Sema3a*<sup>-/-</sup> calvarial cells (Fig. 5a).

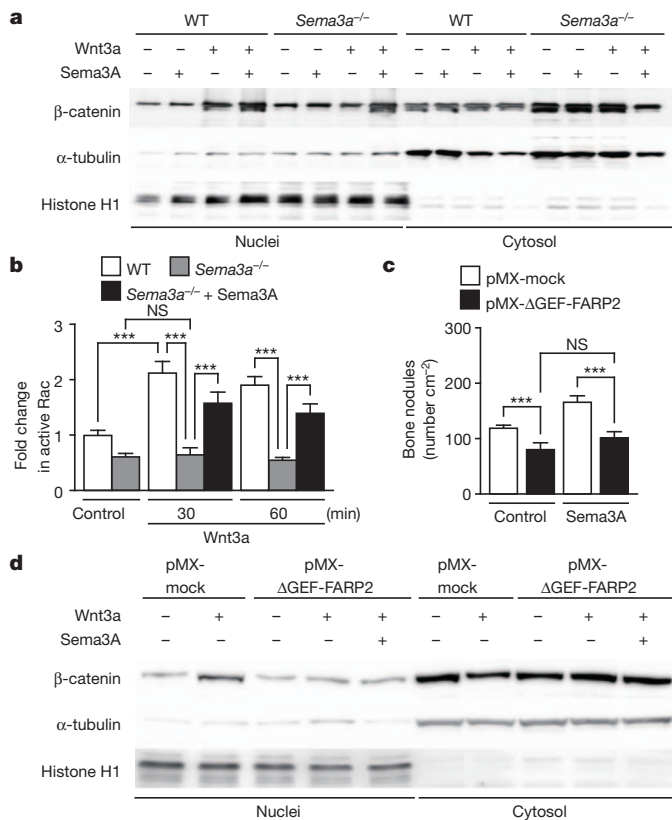
The Sema3A signalling pathway induces activation of the small G protein Rac1 through FARP2 (FERM, RhoGEF and pleckstrin domain protein 2), which is a guanyl-nucleotide exchange factor (GEF) specific for Rac1 (ref. 34). Previous studies suggested a regulatory role of FARP2 in transduction of semaphorin-induced repulsive cues in axons, via Rac1 activation<sup>34</sup>. Because Rac1 promotes the nuclear localization of  $\beta$ -catenin in response to canonical Wnt ligands<sup>35</sup>, we examined the activation of Rac proteins and RhoA in *Sema3a*<sup>-/-</sup> cells. The activation of Rac, but not RhoA, in *Sema3a*<sup>-/-</sup> calvarial cells in response to Wnt3a treatment was significantly decreased (Fig. 5b and Supplementary Fig. 7f) as compared with that in wild-type control cells. In addition, Sema3A treatment facilitated the nuclear translocation of  $\beta$ -catenin and the activation of Rac in *Sema3a*<sup>-/-</sup> cells (Fig. 5a, b). The ectopic expression of a dominant negative form of FARP2 ( $\Delta$ GEF-FARP2)<sup>36</sup> in calvarial cells resulted in the inhibition of osteoblast differentiation, even in the presence of Sema3A (Fig. 5c and Supplementary Fig. 7g). When  $\Delta$ GEF-FARP2 was overexpressed, reduced nuclear accumulation of  $\beta$ -catenin was observed and Sema3A treatment had no effect on its nuclear localization (Fig. 5d). These results indicate that Sema3A stimulates the canonical Wnt/ $\beta$ -catenin signalling pathway, at least in part, through FARP2-mediated activation of Rac1 during osteoblast differentiation.

### Sema3A as an osteoprotective therapeutic agent

To determine the *in vivo* effect of Sema3A administration on bone metabolism, 5-week-old male mice were intravenously injected with recombinant Sema3A or saline once a week. After four weeks of treatment, the trabecular bone volume and trabecular parameters in the distal femur were increased in the Sema3A-treated mice (Fig. 6a and Supplementary Fig. 8a). Bone morphometric analysis showed a decrease in osteoclastic parameters and an increase in osteoblastic parameters (Fig. 6b, c and Supplementary Fig. 8b–e), suggesting that Sema3A exerts a bone-increasing effect by stimulating osteoblastic bone formation and inhibiting osteoclastic bone resorption synchronously. We could not detect any pathological findings in vital organs or any behavioural abnormalities after the Sema3A injection (data not shown). The number of osteoclast precursor cells and osteoprogenitor cells in the bone marrow was not influenced by Sema3A administration (Supplementary Fig. 8f and data not shown), but bone marrow mesenchymal cells derived from Sema3A-treated mice tended to differentiate into osteoblasts instead of adipocytes *in vitro*, although the number of colony forming units was unchanged (Supplementary Fig. 8g, h).

We further investigated the therapeutic potential of Sema3A in a bone regeneration model of cortical bone defects induced by drill hole injury<sup>37</sup>. Microcomputed tomography analysis showed that the regenerated cortical bone volume in Sema3A-treated mice was higher than in saline-treated mice (Fig. 6d and Supplementary Fig. 8i). The significantly increased osteoblast surface and decreased osteoclast surface around the injured region were observed by histomorphometric analysis (Fig. 6e). These results indicate that the local administration of Sema3A into the injured site accelerates bone regeneration, although we cannot rule out the possibility that Sema3A exerted a bone protective effect partly through the regulation of innervation.



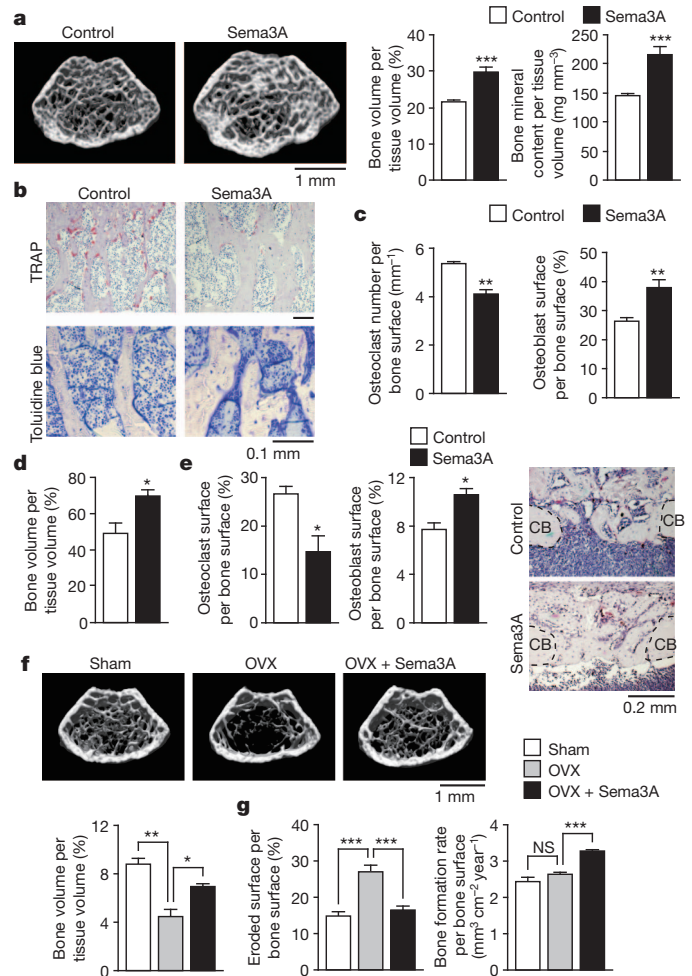


**Figure 5 | Regulation of osteoblast differentiation by Sema3A through canonical Wnt signalling.** **a**, Analysis of nuclear  $\beta$ -catenin levels in the wild-type and *Sema3A*<sup>-/-</sup> calvarial cells stimulated by Wnt3a in the presence or absence of Sema3A. Sema3A was simultaneously added with Wnt3a. **b**, Effect of Sema3A treatment on the activation of Rac in the wild-type and *Sema3A*<sup>-/-</sup> calvarial cells treated with Wnt3a. **c**, Effect of retrovirus-mediated overexpression of  $\Delta$ GEF-FARP2 on bone nodule formation in calvarial cells cultured in the osteogenic medium in the presence or absence of Sema3A. **d**, Effect of retrovirus-mediated overexpression of  $\Delta$ GEF-FARP2 on the nuclear localization of  $\beta$ -catenin in calvarial cells treated with Wnt3a in the presence or absence of Sema3A. Error bars (b, c) denote mean  $\pm$  s.e.m. \*\*\* $P$  < 0.005.

We examined the effect of Sema3A administration on bone loss in an ovariectomized mouse model of postmenopausal osteoporosis. Ovariectomized 9-week-old mice were treated with a weekly intravenous injection of Sema3A starting two days after ovariectomy and continuing for four weeks. Sema3A administration decreased bone loss after ovariectomy both by inhibiting osteoclastic bone resorption and promoting osteoblastic bone formation (Fig. 6f, g and Supplementary Fig. 8j–l). Recombinant human SEMA3A suppressed osteoclastogenesis and promoted osteoblastogenesis in cultured human cells (Supplementary Fig. 9a, b). These results indicate that Sema3A is a promising potential therapeutic target for bone diseases.

## Conclusions

This study demonstrates that the Sema3A expressed by osteoblast lineage cells functions as a potent osteoprotective factor by synchronously inhibiting bone resorption and promoting bone formation (Supplementary Fig. 10a–c). Sema3A represents the long sought soluble molecule with the capacity to bring both osteoblasts and osteoclasts into a condition that favours bone mineral increase. Bone remodelling consists of resorption, transition and formation phases, and the transition phase is under the control of classical coupling factors such as insulin-like growth factor and transforming growth factor- $\beta$ , which link bone resorption with formation<sup>38,39</sup>. Sema3A may have a crucial role in the bone formation phase, in which osteoblasts extensively produce bone, and at the same time restrain osteoclasts from



**Figure 6 | Sema3A as a potential bone-increasing agent.** **a**, Microcomputed tomography analysis of the femurs of 9-week-old wild-type mice treated with Sema3A or saline control ( $n$  = 4–7). **b**, Histological analysis of the proximal tibiae of wild-type mice treated with Sema3A or saline control. **c**, Parameters for osteoclasts and osteoblasts in the bone morphometric analysis of wild-type mice treated with Sema3A or saline control ( $n$  = 4–7). **d**, Microcomputed tomography analysis of bone regeneration of femoral cortex after drill-hole injury ( $n$  = 5). **e**, Histomorphometric analysis of the injured site of the femur (TRAP and haematoxylin staining;  $n$  = 5). CB, cortical bone. **f**, Microcomputed tomography analysis of the femurs of the sham-operated (Sham), ovariectomized (OVX) and Sema3A-treated OVX mice ( $n$  = 4–5). **g**, Parameters for osteoclastic bone resorption and osteoblastic bone formation in the bone morphometric analysis of the Sham, OVX and Sema3A-treated OVX mice ( $n$  = 4–5). Error bars (a, c–g) denote mean  $\pm$  s.e.m. \* $P$  < 0.05; \*\* $P$  < 0.01; \*\*\* $P$  < 0.005.

migrating to the formation sites and starting to resorb the newly formed bone. The Sema3A protein level in the serum or bone microenvironment could be an auspicious biomarker for bone turnover, as we observed that the serum level of Sema3A decreased with age in mice (data not shown).

The potent anti-osteoclastogenic function of Sema3A is tightly controlled by Nrp1 expression regulated by RANKL signalling. Unless Nrp1 is downregulated by RANKL, the Sema3A–Nrp1 axis inhibits osteoclast differentiation by sequestering PlxnA1 from TREM2 so as to suppress ITAM signalling, and also inhibits RhoA activation via Nrp1–PlxnA to suppress osteoclast precursor cell migration (Supplementary Fig. 10b). After RANKL reduces Nrp1 expression, PlxnA1 associates with TREM2 and DAP12, which facilitate the ITAM-mediated calcium signalling required for osteoclast differentiation (Supplementary Fig. 10b). Because PlxnA1, PlxnA2 and PlxnA3, but not PlxnA4, are expressed by osteoclast and osteoblast

lineage cells (data not shown), the relative contribution of these receptor components should be explored in the future.

Therapeutic agents capable of increasing bone formation have essentially been unavailable except for parathyroid hormone or anti-sclerostin antibody<sup>40</sup>. This study may provide a molecular basis for the development of a combined anti-resorptive and bone-increasing agent capable of facilitating bone regeneration.

## METHODS SUMMARY

**Mice and bone analysis.** The generation of *Sema3a*<sup>-/-</sup>, *Nrp1*<sup>Sema-</sup>, *Fos*<sup>-/-</sup>, *Nfatc1*<sup>-/-</sup>, *Tnfrsf11b*<sup>-/-</sup> and *Tyrobp*<sup>-/-</sup> mice was described previously<sup>19,24,41–44</sup>. All mice were maintained under specific pathogen-free conditions. All animal experiments were approved by the Institutional Animal Care and Use Committee of Tokyo Medical and Dental University. Three-dimensional micro-computed tomography analyses and bone morphometric analyses were performed as described<sup>11,14,30,45</sup>. The radiographs were obtained with a high-resolution soft X-ray system (SOFTEX).

**Quantitative RT-PCR analysis and GeneChip analysis.** Real-time quantitative PCR with reverse transcription (RT-PCR) analysis was performed as described<sup>11,14,30,45</sup>. In brief, total RNA was extracted by ISOGEN (NIPPON GENE) according to the manufacturer's instructions. First-strand complementary DNAs were synthesized using Superscript III reverse transcriptase (Invitrogen). Quantitative RT-PCR analysis was performed with the LightCycler apparatus (Roche Applied Science) using SYBR Green Realtime PCR Master Mix (TOYOBO). All primer sequences are available on request. GeneChip analysis and gene set enrichment analysis were performed as described previously<sup>46</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 11 November 2011; accepted 27 February 2012.

Published online 18 April 2012.

1. Takayanagi, H. Osteoimmunology: shared mechanisms and crosstalk between the immune and bone systems. *Nature Rev. Immunol.* **7**, 292–304 (2007).
2. Elefteriou, F. Regulation of bone remodeling by the central and peripheral nervous system. *Arch. Biochem. Biophys.* **473**, 231–236 (2008).
3. Seeman, E. & Delmas, P. D. Bone quality—the material and structural basis of bone strength and fragility. *N. Engl. J. Med.* **354**, 2250–2261 (2006).
4. Teitelbaum, S. L. & Ross, F. P. Genetic regulation of osteoclast development and function. *Nature Rev. Genet.* **4**, 638–649 (2003).
5. Martin, T. J. & Sims, M. Osteoclast-derived activity in the coupling of bone formation to resorption. *Trends Mol. Med.* **11**, 76–81 (2005).
6. Lewiecki, E. M. New targets for intervention in the treatment of postmenopausal osteoporosis. *Nature Rev. Rheumatol.* **7**, 631–638 (2011).
7. Rachner, T. D., Khosla, S. & Hofbauer, L. C. Osteoporosis: now and the future. *Lancet* **377**, 1276–1287 (2011).
8. Reid, I. R. *et al.* Effects of denosumab on bone histomorphometry: the FREEDOM and STAND studies. *J. Bone Miner. Res.* **25**, 2256–2265 (2010).
9. Odvina, C. V. *et al.* Severely suppressed bone turnover: a potential complication of alendronate therapy. *J. Clin. Endocrinol. Metab.* **90**, 1294–1301 (2005).
10. Suda, T. *et al.* Modulation of osteoclast differentiation and function by the new members of the tumor necrosis factor receptor and ligand families. *Endocr. Rev.* **20**, 345–357 (1999).
11. Nakashima, T. *et al.* Evidence for osteocyte regulation of bone homeostasis through RANKL expression. *Nature Med.* **17**, 1231–1234 (2011).
12. Xiong, J. *et al.* Matrix-embedded cells control osteoclast formation. *Nature Med.* **17**, 1235–1241 (2011).
13. Simonet, W. S. *et al.* Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell* **89**, 309–319 (1997).
14. Luo, Y., Raible, D. & Raper, J. A. Collapse in brain that induces the collapse and paralysis of neuronal growth cones. *Cell* **75**, 217–227 (1993).
15. Tran, T. S., Kolodkin, A. L. & Bharadwaj, R. Semaphorin regulation of cellular morphology. *Annu. Rev. Cell Dev. Biol.* **23**, 263–292 (2007).
16. Negishi-Koga, T. *et al.* Suppression of bone formation by osteoclastic expression of semaphorin 4D. *Nature Med.* **17**, 1473–1480 (2011).
17. Takegahara, N. *et al.* Plexin-A1 and its interaction with DAP12 in immune responses and bone homeostasis. *Nature Cell Biol.* **8**, 615–622 (2006).
18. Matsuo, K. & Irie, N. Osteoclast-osteoblast communication. *Arch. Biochem. Biophys.* **473**, 201–209 (2008).
19. Taniguchi, M. *et al.* Disruption of *semaphorin III/D* gene causes severe abnormality in peripheral nerve projection. *Neuron* **19**, 519–530 (1997).
20. Gomez, C. *et al.* Expression of Semaphorin-3A and its receptors in endochondral ossification: potential role in skeletal development and innervation. *Dev. Dyn.* **234**, 393–403 (2005).
21. Behar, O., Golden, J. A., Mashimo, H., Schoen, F. J. & Fishman, M. C. Semaphorin III is needed for normal patterning and growth of nerves, bones and heart. *Nature* **383**, 525–528 (1996).
22. Jacquin, C., Gran, D. E., Lee, S. K., Lorenzo, J. A. & Aguila, H. L. Identification of multiple osteoclast precursor populations in murine bone marrow. *J. Bone Miner. Res.* **21**, 67–77 (2006).
23. Neufeld, G. & Kessler, O. The semaphorins: versatile regulators of tumour progression and tumour angiogenesis. *Nature Rev. Cancer* **8**, 632–645 (2008).
24. Gu, C. *et al.* Neuropilin-1 conveys semaphorin and VEGF signaling during neural and cardiovascular development. *Dev. Cell* **5**, 45–57 (2003).
25. Ashburner, B. P., Westerheide, S. D. & Baldwin, A. S. Jr. The p65 (RelA) subunit of NF- $\kappa$ B interacts with the histone deacetylase (HDAC) corepressors HDAC1 and HDAC2 to negatively regulate gene expression. *Mol. Cell Biol.* **21**, 7065–7077 (2001).
26. Takayanagi, H. *et al.* Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts. *Dev. Cell* **3**, 889–901 (2002).
27. Koga, T. *et al.* Costimulatory signals mediated by the ITAM motif cooperate with RANKL for bone homeostasis. *Nature* **428**, 758–763 (2004).
28. Takahashi, T. & Strittmatter, S. M. PlexinA1 autoinhibition by the plexin sema domain. *Neuron* **29**, 429–439 (2001).
29. Narazaki, M. & Tosato, G. Ligand-induced internalization selects use of common receptor neuropilin-1 by VEGF165 and semaphorin3A. *Blood* **107**, 3892–3901 (2006).
30. Nishikawa, K. *et al.* Maf promotes osteoblast differentiation in mice by mediating the age-related switch in mesenchymal cell differentiation. *J. Clin. Invest.* **120**, 3455–3465 (2010).
31. Gimble, J. M., Zvonice, S., Floyd, Z. E., Kassem, M. & Nuttall, M. E. Playing with bone and fat. *J. Cell. Biochem.* **98**, 251–266 (2006).
32. Krishnan, V., Bryant, H. U. & MacDougald, O. A. Regulation of bone mass by Wnt signaling. *J. Clin. Invest.* **116**, 1202–1209 (2006).
33. Takada, I., Kouzmenko, A. P. & Kato, S. Wnt and PPAR $\gamma$  signaling in osteoblastogenesis and adipogenesis. *Nature Rev. Rheumatol.* **5**, 442–447 (2009).
34. Toyofuku, T. *et al.* FARP2 triggers signals for Sema3A-mediated axonal repulsion. *Nature Neurosci.* **8**, 1712–1719 (2005).
35. Wu, X. *et al.* Rac1 activation controls nuclear localization of  $\beta$ -catenin during canonical Wnt signaling. *Cell* **133**, 340–353 (2008).
36. Takegahara, N. *et al.* Integral roles of a guanine nucleotide exchange factor, FARP2, in osteoclast podosome rearrangements. *FASEB J.* **24**, 4782–4792 (2010).
37. Nagashima, M. *et al.* Bisphosphonate (YM529) delays the repair of cortical bone defect after drill-hole injury by reducing terminal differentiation of osteoblasts in the mouse femur. *Bone* **36**, 502–511 (2005).
38. Tang, Y. *et al.* TGF- $\beta$ 1-induced migration of bone mesenchymal stem cells couples bone resorption with formation. *Nature Med.* **15**, 757–765 (2009).
39. Hayden, J. M., Mohan, S. & Baylink, D. J. The insulin-like growth factor system and the coupling of formation to resorption. *Bone* **17**, S93–S98 (1995).
40. Kawai, M., Mödder, U. I., Khosla, S. & Rosen, C. J. Emerging therapeutic opportunities for skeletal restoration. *Nature Rev. Drug Discov.* **10**, 141–156 (2011).
41. Grigoriadis, A. E. *et al.* c-Fos: a key regulator of osteoclast-macrophage lineage determination and bone remodeling. *Science* **266**, 443–448 (1994).
42. Asagiri, M. *et al.* Autoamplification of NFATc1 expression determines its essential role in bone homeostasis. *J. Exp. Med.* **202**, 1261–1269 (2005).
43. Mizuno, A. *et al.* Severe osteoporosis in mice lacking osteoclastogenesis inhibitory factor/osteoprotegerin. *Biochem. Biophys. Res. Commun.* **247**, 610–615 (1998).
44. Kaifu, T. *et al.* Osteopetrosis and thalamic hypomyelination with synaptic degeneration in DAP12-deficient mice. *J. Clin. Invest.* **111**, 323–332 (2003).
45. Hayashi, M. *et al.* Ly49Q, an ITIM-bearing NK receptor, positively regulates osteoclast differentiation. *Biochem. Biophys. Res. Commun.* **393**, 432–438 (2010).
46. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to D. D. Ginty and A. L. Kolodkin for providing the *Nrp1*<sup>Sema-</sup> knock-in mice. We thank Y. Goshima for providing vectors and technical help. We thank A. Yamaguchi, H. Asahara and F. Suto for providing reagents and technical help. We also thank K. Okamoto, T. Negishi-Koga, K. Nishikawa, H. Inoue, T. Suda, T. Ando, Y. Kunisawa, Y. Ogihara and S. Fukuse for discussion and assistance. This work was supported in part by a grant for the Exploratory Research for Advanced Technology Program, the Takayanagi Osteonetwork Project from the Japan Science and Technology Agency; Grant-in-Aid for Young Scientist A from the Japan Society for the Promotion of Science (JSPS); a Grant-in-Aid for Challenging Exploratory Research from the JSPS; grants for the Global Center of Excellence Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan; and grants from the Tokyo Biochemical Research Foundation, the Life Science Foundation of Japan, Takeda Science Foundation, Uehara Memorial Foundation, Naito Foundation, BMKK RA Research Fund and Astellas Foundation for Research on Metabolic Disorders.

**Author Contributions** M.H. performed most of the experiments, interpreted the results and prepared the manuscript. T.N. performed immunohistochemical experiments and provided advice on project planning and data interpretation and prepared the manuscript. M.T. provided technical help. T.K. conducted the GeneChip analysis. A.K. provided advice on project planning and technical help. H.T. directed, supervised the project and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to H.T. ([taka.csi@tmd.ac.jp](mailto:taka.csi@tmd.ac.jp)).



## METHODS

**Cell culture.** For *in vitro* osteoclast differentiation in the monoculture system, primary bone marrow cells ( $1 \times 10^5$  cells per  $\text{cm}^2$ ) were suspended in culture medium ( $\alpha$ -MEM containing penicillin, streptomycin and 10% FBS) supplemented with  $10 \text{ ng ml}^{-1}$  M-CSF (R&D Systems) for two days to obtain BMMs. The resultant BMMs were further cultured in medium supplemented with  $10 \text{ ng ml}^{-1}$  M-CSF and  $5\text{--}50 \text{ ng ml}^{-1}$  RANKL (PeproTech) for three days. Culture medium was changed every second day. Where indicated, calvarial cell-conditioned medium with or without soluble Nrp1 (R&D Systems) was added 12 h before the RANKL stimulation. Sema3A-Fc (R&D Systems) was added 12 h before or 12 h after the RANKL stimulation. For Sema6D treatment, BMMs were collected and seeded onto culture plates coated with soluble recombinant Sema6D (R&D Systems), and cultured with  $10 \text{ ng ml}^{-1}$  M-CSF and  $5 \text{ ng ml}^{-1}$  RANKL for three days. For the generation of osteoclast *in vitro* in the coculture system, primary bone marrow cells ( $5 \times 10^4$  cells per  $\text{cm}^2$ ) and calvarial cells ( $5 \times 10^3$  cells per  $\text{cm}^2$ ) were cultured in the presence of  $10 \text{ nM}$   $1\alpha,25\text{-dihydroxyvitamin D}_3$  and  $1 \mu\text{M}$  prostaglandin E2 for 4–6 days. For human osteoclast differentiation, human peripheral blood mononuclear cells were separated from peripheral blood obtained from healthy volunteers by density gradient centrifugation with Lymphoprep (AXIS-SHIELD). Cells ( $2 \times 10^5$  cells per 0.5 ml) were cultured in  $\alpha$ -MEM with 10% FBS supplemented with  $30 \text{ ng ml}^{-1}$  M-CSF for two days. The resultant preosteoclasts were further cultured in medium supplemented with  $30 \text{ ng ml}^{-1}$  M-CSF and  $60 \text{ ng ml}^{-1}$  RANKL for four days. Culture medium was changed every second day. The differentiation of osteoclasts was evaluated by TRAP staining. A NF- $\kappa$ B activation inhibitor (6-amino-4-(phenoxyphenylethylamino)quinazoline; Calbiochem) was used to inhibit NF- $\kappa$ B activity. The concentration of intracellular calcium was measured as described<sup>27</sup>. For *in vitro* osteoblast differentiation, calvarial cells were isolated from the calvarial bone of newborn mice by enzymatic digestion in  $\alpha$ -MEM with 0.1% collagenase and 0.2% dispase, and were cultured with  $\alpha$ -MEM with 10% FBS. After two days, cells were reseeded ( $1 \times 10^4$  cells per  $\text{cm}^2$ ) and cultured with osteogenic medium (100 mM ascorbic acid, 5 mM  $\beta$ -glycerophosphate and 10 nM dexamethasone). Culture medium was changed every third day. After seven days, ALP staining and activity measurement were performed, and after 21 days, bone nodule formation was assessed by alizarin red staining. Human mesenchymal stem cells (Lonza) were cultured according to the manufacturer's protocol. To induce adipocyte differentiation *in vitro*, primary bone marrow cells ( $5 \times 10^5$  cells per  $\text{cm}^2$ ) were cultured with  $\alpha$ -MEM containing 10% FBS. After 24 h, non-adherent cells were removed and adherent cells were cultured with adipogenic medium (0.5 mM 3-isobutyl-1-methylxanthine,  $5 \mu\text{g ml}^{-1}$  insulin and  $1 \mu\text{M}$  dexamethasone) for 10–14 days. Culture medium was changed every second day. Lipid accumulation in adipocytes was determined with Oil Red O staining. Cell proliferation was determined using a cell proliferation ELISA kit (Roche Applied Science). The percentage of apoptotic cells was determined by TUNEL (TdT-mediated dUTP nick end labelling) staining with the MEBSTAIN apoptosis kit direct (MBL). For colony forming unit (CFU) assays, primary bone marrow cells ( $2.5 \times 10^5$  cells per  $\text{cm}^2$ ) were seeded and cultured with MesenCult basal medium supplemented with mesenchymal stem cell stimulatory supplements (StemCell Technologies). On day 10, the cells were stained with toluidine blue. For CFU-ALP and CFU-osteoblast (CFU-OB) assays, primary bone marrow cells ( $2.5 \times 10^5$  cells per  $\text{cm}^2$ ) were seeded and cultured with MesenCult basal medium and mesenchymal stem cell stimulatory supplements plus 100 mM ascorbic acid, 5 mM  $\beta$ -glycerophosphate and 10 nM dexamethasone. On day 10, CFU-ALP colonies were stained for ALP. Mineral deposition was determined with von Kossa staining of CFU-OB colonies on day 25. Primary osteoblasts and osteocytes were isolated from the long bones of CAG-CAT-EGFP/Dmp1-Cre double transgenic mice as described<sup>11</sup>.

**Purification and identification of inhibitory factor of osteoclast differentiation.** Mouse calvarial cells were statically cultured with  $\alpha$ -MEM supplemented with 1% FBS. Medium was conditioned for 72 h and filtered to remove non-adherent cells and debris. Conditioned medium was concentrated by ammonium sulphate precipitation (40% saturation), and the pellet was suspended in 10 mM sodium phosphate buffer, pH 7.4, and desalted using a PD-10 column (GE Healthcare). Concentrated conditioned media were then loaded onto a Mono Q 5/50 column (GE Healthcare) in 10 mM sodium phosphate buffer, pH 7.4, and proteins binding to the Mono Q matrix were eluted by a gradient of 0–100% 1 M NaCl and 10 mM sodium phosphate buffer, pH 7.4. Protein concentrations were determined by the absorbance at 280 nm. The effect of the fractionated conditioned media on osteoclast differentiation was examined by directly adding each fraction to RANKL-induced osteoclast differentiation *in vitro*. Fractions 33–35 contained a high concentration of NaCl, which exerts an inhibitory effect on osteoclast differentiation. Proteins of the highly inhibitory fractions were dissolved in SDS-PAGE sample buffer (Nacalai Tesque) and the sample was resolved by SDS-PAGE. Protein bands were visualized by Coomassie brilliant blue staining and all the protein bands were excised by scalpel. The samples were

analysed using nano-LC-MS/MS by Japan Bioservice. The data were submitted to the MASCOT program for identification.

**Immunohistochemical staining.** After fixation in 4% paraformaldehyde, bone tissues were decalcified in 10% EDTA at 4 °C for 2 weeks and embedded in paraffin after dehydration. For immunohistochemical staining, antigen retrieval was carried out with 10 mM citric acid, pH 6.0, at room temperature for 2 h. After quenching of endogenous peroxidase activity by incubation with 3%  $\text{H}_2\text{O}_2$  in methanol, the sections were incubated with an anti-Sema3A polyclonal antibody (Santa Cruz Biotechnology) in immunoreaction enhancer solution (Can Get Signal immunostain, TOYOBO) at 4 °C for overnight. After washing with PBS, the sections were incubated with peroxidase-conjugated secondary antibody according to the manufacturer's instructions (histofine, Nichirei Bioscience). The signals were visualized with 3,3'-diaminobenzidine tetrahydrochloride and  $\text{H}_2\text{O}_2$ . TRAP staining was conducted after the immunostaining. Haematoxylin was used for nuclear counterstaining.

**Western blot and immunoprecipitation analyses.** Cell lysate or culture supernatant of calvarial cells was subjected to western blot analysis using the specific antibodies for Nrp1 (Calbiochem),  $\beta$ -actin (Sigma-Aldrich), Sema3A, Nrp1, p50, p65, histone H1 (Santa Cruz Biotechnology), phospho-ERK, ERK, phospho-JNK, JNK, phospho-p38, p38, phospho-IKK $\alpha/\beta$ , IKK $\alpha$ , IKK $\beta$ , phospho-PLC $\gamma$ 2, PLC $\gamma$ 2 (Cell Signaling Technology),  $\beta$ -catenin (Millipore) and  $\alpha$ -tubulin (MBL). Nuclear proteins were prepared with nuclear extract kit in accordance with the manufacturer's protocol (Active Motif). For immunoprecipitation analysis, cells were solubilized in lysis buffer (1% Nonidet P-40 in 50 mM NaCl, 50 mM Tris-HCl, 5 mM EDTA, 1 mM NaF and 2 mM PMSF), supplemented with complete protease inhibitor cocktail (Roche Applied Science). Immunoprecipitation was performed by incubation with an anti-Flag M2 (Sigma-Aldrich) or anti-PlxnA1 antibody (Santa Cruz Biotechnology) followed by the addition of dynabeads protein G (Invitrogen). Immune complexes were separated by electrophoresis followed by blotting with anti-Flag M2, anti-V5 (Invitrogen), anti-Nrp1, anti-PlxnA1, anti-DAP12 (Santa Cruz Biotechnology) and anti-TREM2 antibodies (R&D Systems).

**Flow cytometric analysis.** For the analysis of bone marrow-derived osteoclast precursor cells, a single cell suspension of mouse bone marrow cells was stained with anti-CD3e (145-2C11, eBioscience), anti-B220 (RA3-6B2, eBioscience), anti-CD11b (M1/70, eBioscience), anti-CD115 (AFS98, eBioscience) and anti-CD117 (2B8, eBioscience) antibodies. For intracellular staining of BMMs, anti-CD11b, anti-CD115 and anti-Nrp1 (R&D Systems) were used. Flow cytometric analysis was performed using FACSCantoII with Diva software (BD Biosciences).

**ELISA.** Concentrations of soluble RANKL and Opg in serum were determined using ELISA kits (R&D Systems), according to the manufacturer's instruction.

**Retroviral gene transfer.** The retroviral vector pMXs-Nrp1-IRES-EGFP was constructed by inserting DNA fragments encoding Nrp1 into pMXs-IRES-EGFP. The construction of the retroviral vectors pMX-FARP2-IRES-GFP and pMX-AGEF-FARP2-IRES-GFP was described previously<sup>36</sup>. For the construction of the retroviral vectors pSIREN-RetroQ-ZsGreen-shNrp1 and pSIREN-RetroQ-ZsGreen-shControl, RNA targeting regions with a hairpin sequence (Nrp1 shRNA sense: 5'-GCCCGAATGTTCTCAGAACTACTCGAGTAGTCTGAGAACATTCGGGCTTTT-3'; Nrp1 shRNA antisense: 5'-AAAAAGCCCCGATGTTCTCAGAACTACTCGAGTAGTCTGAGAACATTCGGGCT-3'; control shRNA sense: 5'-GTGCGTTGCTAGTACCACTTCAAGAGATTTTACGCGT-3'; control shRNA antisense: 5'-ACGCGTAAAAATCTCTTGAAGTTGGTACTAGCAACGAC-3') were inserted into RNAi-ready pSIREN-RetroQ-ZsGreen (Clontech). The retrovirus supernatants were obtained by transfecting the retroviral vectors into the Plat-E packaging cell line using FuGENE 6 (Roche Applied Science).

**Chromatin immunoprecipitation assay.** Chromatin immunoprecipitation assay was performed using the ChIP-IT express chromatin immunoprecipitation kit (Active Motif) according to the manufacturer's instructions. The antibodies used for immunoprecipitation were anti-p50, anti-p65 and normal rabbit IgG (Santa Cruz Biotechnology). The primer sequences were as follows: Nrp1 region 1, 5'-CATACGTGACCTTGGCTCT-3' and 5'-CCTGGCTGGAGATTCAGA GA-3'; Nrp1 region 2, 5'-ACCTTACCCACAGCTCCTT-3' and 5'-ATACGCCACCACTTACGAG-3'; Nrp1 region 3, 5'-ATGTGGCTTGGTGAAAG GAG-3' 5'-TGCTTCTACCTTCGGGTGAT-3'.

**Reporter gene assay.** The reporter plasmid Nrp1-Luc was constructed by subcloning a 3,259 base pair fragment of the 5' flanking region of the mouse Nrp1 gene into the pGL3-basic vector (Promega). The reporter plasmids and the expression plasmids were transfected into NIH3T3 cells using FuGENE 6 (Roche Applied Science). After 36 h, dual luciferase assay was performed according to the manufacturer's protocol (Promega).

**Migration assay.** BMMs suspended in complete medium were added to the upper chamber of transwell units (Corning). Inserts were placed into the lower chambers



of transwell units containing M-CSF with or without Sema3A-Fc. After incubation, cells were fixed with 4% paraformaldehyde and stained with 0.5% toluidine blue. The cells on the upper side of the membrane were removed and the cells that had migrated to the lower side of the membrane and chamber were counted.

**G-LISA small G protein activation assay.** RhoA and Rac GTPase activation were determined using the G-LISA RhoA and Rac absorbance-based assay (Cytoskeleton) according to the manufacturer's instructions. In brief, cell lysates were prepared and normalized. After the addition of antibodies against RhoA or Rac and the incubation with the horseradish peroxidase detection reagent, signals were detected with a spectrophotometer.

**In vivo treatment with recombinant Sema3A.** Five-week-old C57BL/6 mice were given weekly intravenous injections of 1 mg per kg body weight of Sema3A-Fc or vehicle for four weeks. Three days after the last injection, bone analysis was performed as described earlier.

**Bone regeneration model.** Skeletal injury was generated as described previously<sup>37</sup>. In brief, C57BL/6 mice were anaesthetized with an intraperitoneal

injection of pentobarbital sodium. A 5-mm longitudinal incision was made over the proximal femur and the bone surface was exposed by splitting the muscle. A 0.5-mm hole was made by drilling through the anterior portion of the diaphysis of the bilateral femurs. After four and seven days of surgery, femoral defects were treated with Sema3A-Fc (0.5 mg per kg body weight) by injection into the injury site. Mice were euthanized at day 14 after surgery and bone analyses were performed.

**Ovariectomy-induced bone loss.** Nine-week-old female mice were ovariectomized or sham operated. More than five mice were examined in each group. Ovariectomized mice were given weekly intravenous injections of 1 mg per kg body weight of Sema3A-Fc or vehicle for four weeks. Three days after the last injection, all of the mice were euthanized and subjected to bone analysis as described earlier.

**Statistical analyses.** Statistical analyses were performed using the unpaired two-tailed Student's *t* test (\**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; NS, not significant, throughout the paper). All data are expressed as the mean ± s.e.m. Results are representative examples of more than three independent experiments.

# Impact spherules as a record of an ancient heavy bombardment of Earth

B. C. Johnson<sup>1</sup> & H. J. Melosh<sup>1,2</sup>

**Impact craters are the most obvious indication of asteroid impacts, but craters on Earth are quickly obscured or destroyed by surface weathering and tectonic processes<sup>1</sup>. Earth's impact history is inferred therefore either from estimates of the present-day impactor flux as determined by observations of near-Earth asteroids, or from the Moon's incomplete impact chronology<sup>2–4</sup>. Asteroids hitting Earth typically vaporize a mass of target rock comparable to the projectile's mass. As this vapour expands in a large plume or fireball, it cools and condenses into molten droplets called spherules<sup>5</sup>. For asteroids larger than about ten kilometres in diameter, these spherules are deposited in a global layer. Spherule layers preserved in the geologic record accordingly provide information about an impact even when the source crater cannot be found<sup>1</sup>. Here we report estimates of the sizes and impact velocities of the asteroids that created global spherule layers. The impact chronology from these spherule layers reveals that the impactor flux was significantly higher 3.5 billion years ago than it is now. This conclusion is consistent with a gradual decline of the impactor flux after the Late Heavy Bombardment.**

There have been several attempts to model the process of spherule formation in the hope that the properties of an impacting body could be determined from observations of the resulting spherule layer. These simplified models indicated that spherule size depends strongly on the size of an impactor but suggested a weak dependence on the impact velocity<sup>6–8</sup>. A more detailed model contradicts these results and instead shows that the impact velocity, not the size of an impactor, is the main determinant of spherule size<sup>5</sup>. Although we cannot use spherule size to deduce the size of an impactor, we can use the thickness of a spherule layer to do so. We assume that, as with the Cretaceous–Palaeogene boundary layer, all spherule layers have a roughly constant thickness globally<sup>9</sup>. This means the total mass of spherules in the layer can be determined using only the thickness of the spherule layer and the fraction of spherules in the layer at one, or a few, representative

locations. In the Supplementary Information, we derive an equation that relates impactor size to spherule layer thickness:

$$D_{\text{imp}} = 17(t_r/\xi)^{(1/3)} \quad (1)$$

where  $D_{\text{imp}}$  is the impactor diameter in kilometres and  $\xi$  is an efficiency factor that conservatively ranges from 0.5 to 2 for typical asteroidal impact velocities on Earth. Additionally,  $t_r$  is the layer's reduced thickness in centimetres, defined as  $t_r = 2f_{\text{sp}}t$  where  $t$  is the measured layer thickness and  $f_{\text{sp}}$  is the volume fraction of spherules in the layer.

We test the accuracy of equation (1) by comparing the impactor size estimated using spherule layer thickness with the impactor size determined by other methods. The Cretaceous–Palaeogene boundary layer is found at numerous sites globally and has a thickness of around 3 mm and consists of about half spherules by volume<sup>9</sup>. Using the entire range of  $\xi$  we find  $D_{\text{imp}} = 9.0\text{--}14$  km. This is consistent with the size of the Chicxulub impactor ( $10 \pm 4$  km) as determined by iridium fluence and similar estimates from the size of the Chicxulub impact structure<sup>10,11</sup>. Our impactor size estimates for two other spherule layers, S2 and S3 (Table 1), are also consistent with the estimate, 3–7 times larger than the Chicxulub impactor, which is also based on total iridium fluence<sup>12</sup>.

The estimates of impactor sizes using equation (1) are only valid if the data taken from a limited region are representative of the global spherule layer. Many spherule layers show signs of redeposition by surface processes or subsequent tectonic deformation, which both reduce the accuracy of the global thickness estimate<sup>13</sup>. In addition to this uncertainty, the assumption that these layers represent a global layer of vapour condensates may be incorrect. Some of these spherule beds may be composed of melt spherules created through the fragmentation of melt ejected during an impact. Melt spherules can be morphologically identical to vapour-condensate spherules and may be ejected thousands of kilometres from the impact site<sup>9</sup>. One major difference between the two types of spherules is the dependence of

**Table 1 | Earth's impact history from spherule layer data**

Name of spherule layer	Age of layer (Gyr)	$t_r$ (cm)	Average spherule size (mm)	Impactor diameter (km)	Impact velocity (km s <sup>-1</sup> )	Extraterrestrial origin	Reference
S1	3.47	10–15	0.3–0.6	29–53	18.8–21.2		1, 13
S2	3.26	20	0.15–2.5	37–58	17.7–25.6	Cr	1, 13
S3	3.24	30–35	0.6–1.5	41–70	20.6–22.8	Ir, Cr	13
S4	3.24	15	0.2–1.6	33–53	18.2–22.2	Ir, Cr	1, 13
Jeerinah*	2.63	0.1–0.5	0.5	6.3–17	21.9–25.1	Ir, Cr	19
Monteville*	2.60–2.65	10	0.65	29–46	20.4–21.4	Ir	1, 20
Reivilo†	2.56	2	~0.6	17–27	21.3–22.4	Ir	1, 21
Paraburdoo†	2.57	2	~0.6	17–27	21.3–22.4	Ir	21, 22
Bee Gorge	2.54	0.1–1	0.56	6.3–21	21.7–26.1	Ir	23, 24
Dales Gorge‡	2.49	12	0.6–0.8	31–49	20.1–21.7	Ir, Cr	1, 23, 25
Kuruman‡	2.46–2.52	0.6	~0.6	11–18	22.3–23.8		26
Grænsesø	1.85–2.13	40	0.5–1.0	46–73	19.1–21.3		1, 27
Cretaceous–Palaeogene	0.065	0.3	0.25	9.0–14	20.4–21.5	Ir, Cr	9, 28
Clinopyroxene	0.035	0.04	0.25–0.5	4.6–7.3	22.0–27.0	Ir, Cr	1, 29, 30

\*, † or ‡ Layers with the same symbol are attributed to the same impact. The size of the impactor is as determined by equation (1), and the impact velocity was determined by the method described in the Supplementary Fig. 1. Extraterrestrial origin is indicated by Ir and/or Cr where Ir means that the layer has a significant iridium anomaly and Cr means that chromium isotope data indicates an extraterrestrial origin. There are many other spherule layers that are known to be more proximal melt droplet layers that are not included in this table. Where possible, we obtained spherule layer data from the 2004 review<sup>1</sup> of known spherule layers in an attempt to present a consistent interpretation of geologic data.

<sup>1</sup>Department of Physics, Purdue University, 525 Northwestern Avenue, West Lafayette, Indiana 47907, USA. <sup>2</sup>Department of Earth and Atmospheric Sciences, Purdue University, 550 Stadium Mall Drive, West Lafayette, Indiana 47907, USA.



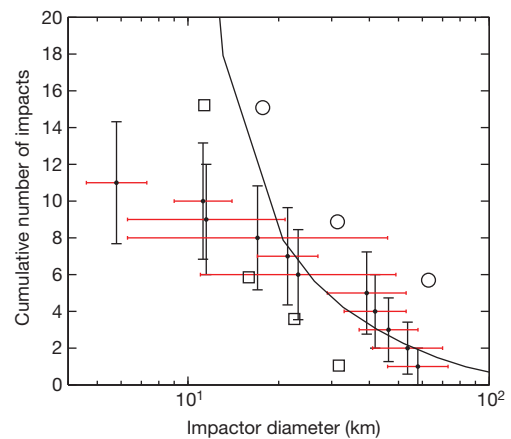
the layer thickness and size of spheroidal particles as a function of distance from the point of impact. Both droplet abundance and particle size decrease strongly with distance from the impact for melt droplets, whereas for a vapour-condensate spherule layer these properties are more or less uniform globally<sup>9</sup>. This means that if a layer is found in several locations, it may be possible to distinguish between the two types of spherules. It may also be possible to determine whether a layer is composed of melt droplets if impact lapilli are present. These accretionary particles are found in association with melt droplets in the Chicxulub ejecta more than 2,000 km from the source crater<sup>14</sup>.

Another way to distinguish between spherule types is to determine whether a layer contains extraterrestrial material<sup>5</sup>. Only vapour-condensate spherules should contain significant extraterrestrial material, because the impactor is vaporized at typical Earth impact velocities. Although chromium isotope data gives the most robust indication of extraterrestrial origin, anomalous iridium and platinum group element content also indicate an extraterrestrial origin<sup>10,12</sup>. It has also been suggested that mineral replacement over billions of years may erase any sign of extraterrestrial material originally present in the spherule layer<sup>1</sup>. For this reason it is possible to confirm a layer's extraterrestrial origin but it is often impossible to rule out such an origin. In the case of Chicxulub, at locations where ejecta containing melt droplets and the ejecta composed of vapour-condensate spherules are both present, the two species are stratigraphically distinct, making it possible to distinguish between vapour-condensate layers and melt droplet layers using anomalous iridium. Owing to redeposition by surface processes, it is possible that other layers will be composed of a mix of melt droplets and vapour-condensate spherules.

In addition to impactor size, the impact velocity is computed from the average spherule size using the model of ref. 5. This method is described in Supplementary Fig. 1. The average estimated impact velocity from all the known spherule layers is about  $21.8 \pm 2.2 \text{ km s}^{-1}$ , close to the expected average of about  $20.3 \text{ km s}^{-1}$  (ref.<sup>15</sup>). The consistency of our impact velocity estimates with the expected average is another indication that the layers we include are indeed vapour-condensate spherule layers. We expect that if the layers were composed of melt droplets, which are made by a completely different process and can be much larger than vapour-condensate spherules, then our velocity estimates would deviate significantly from this average.

The data from Table 1 allows us to construct an impactor size frequency distribution (SFD) for Earth (Fig. 1). We expect that Earth's impactor SFD obtained from spherule layers should exhibit a power-law dependence similar to the impactor SFDs obtained from direct astronomical observations of near-Earth asteroids. Comparing the form of the different impactor SFDs in Fig. 1, we see an obvious deficit of impactors smaller than about 20 km in diameter in the impactor SFD obtained from spherule layer data. It is plausible that this rollover at smaller impactor sizes is due to either observational or preservation biases. As an impacting body becomes smaller, it creates progressively thinner and sparser layers that are more difficult to recognize and are more easily obscured or destroyed by subsequent geologic processes or diluted during initial emplacement<sup>1</sup>.

In addition to a bias in favour of thicker layers, the history of large impacts on Earth provided by spherule layers is probably incomplete, given that only two regions have been systematically searched for Precambrian spherule layers: the Pilbara Craton in Western Australia (especially the Warrawoona and Hamersley successions) and the Kaapvaal Craton in South Africa (especially the Barberton Greenstone belt and Griqualand West succession). It is thus not surprising that most of the known layers were discovered in these two regions. These two regions represent strata from 2.63–2.49 billion years (Gyr) ago and 3.47–3.24 Gyr ago. These strata thus represent approximately ten per cent of Earth's impact history over 3.5 Gyr. Considering the poor sampling of much of the geologic record, it is likely that many more spherule layers are yet to be discovered. Additionally, a preliminary report shows evidence of three more spherule layers in the Barberton Greenstone belt<sup>16</sup>.



**Figure 1 | Earth's impactor SFD.** Earth's impactor SFD is plotted as the cumulative number of impacts larger than a given diameter as a function of diameter. The points with vertical black error bars (s.d.) represent the SFD based on spherule layer data from Table 1, where impactor diameter is reported as a size range. Each black dot represents a single impact, plotted at the geometric mean of this size range, and the horizontal red error bars are plotted to represent the entire impactor size range. The squares, solid curve and circles represent three separate estimates of Earth's cumulative impactor SFD, obtained by multiplying the present-day impactor flux reported as probability of impacts per year by the time interval of 3.5 Gyr, which is approximately the age of bed S1. The squares are from direct observations of near-Earth asteroids<sup>4</sup>. The solid curve is obtained by scaling the SFD of the main asteroid belt to observations of near-Earth asteroids at smaller sizes, where there are many more objects and therefore much less statistical error<sup>2</sup>. The circles are obtained by scaling the SFD of Mars-crossing asteroids to observations of near-Earth asteroids in a similar manner to the solid curve<sup>3</sup>. When the literature indicates that multiple layers come from the same impact, we assume this conclusion is correct. The red error bar then represents the range from the smallest size indicated by the multiple layers to the largest size. For example, we assume that the two spherule beds Jeerinah and Monteville were created by a single impactor which may be anywhere from 6.3 km to 46 km in diameter. The large size range for this particular impact may indicate that one of these layers is from a separate impact or represents a melt droplet layer.

During the Late Heavy Bombardment (LHB), 4.1–3.8 Gyr ago, the impactor flux was much higher than the present-day flux<sup>17</sup>. It is debated whether the post-LHB impactor flux dropped quickly to present-day values or slowly decreased as implied by a Solar System model that includes an extended asteroid belt<sup>18</sup>. The impactor SFDs inferred from observations of near-Earth asteroids represent the expected impactor SFD on Earth, assuming that the impactor flux has remained constant over the past 3.5 Gyr, which is consistent with a quick decrease in post-LHB impactor flux. The impactor SFD implied by spherule layers makes no assumption about the time dependence of the impactor flux and therefore, if complete, represents the average impactor flux over 3.5 Gyr. The spherule record is not complete, representing only about 10% of the total 3.5-Gyr history. Thus, if the impactor flux implied by spherules has remained constant over 3.5 Gyr, we must multiply the current number of impacts by a factor of about ten. This SFD would then disagree with even the highest estimates of current-day impactor flux. Therefore, a constant impactor flux cannot be simultaneously consistent with spherule layer data and observations of near-Earth asteroids. This implies that the impactor flux was significantly higher 3.5–2.5 Gyr ago, consistent with a gradual decline of the post-LHB impactor flux. However, the spherule record cannot rule out other explanations for the heightened impactor flux, including the unlikely occurrence of one or more large spikes in the impactor flux that happen to coincide with the two time frames for which there are well-preserved strata. To make a more robust and quantitative conclusion about the time dependence of the post-LHB impactor flux, a more complete search for spherule layers in the geologic record is required.

Received 18 November 2011; accepted 21 February 2012.

Published online 25 April 2012.

- Simonson, B. M. & Glass, B. P. Spherule layers—records of ancient impacts. *Annu. Rev. Earth Planet. Sci.* **32**, 329–361 (2004).
- Ivanov, B. A. & Hartmann, W. K. In *Treatise on Geophysics* Vol. 10 *Planets and Moons* (ed. Schubert, G.), 202–242 (Elsevier, 2007).
- Le Feuvre, M. & Wieczorek, M. A. Nonuniform cratering of the Moon and a revised crater chronology of the inner Solar System. *Icarus* **214**, 1–20 (2011).
- Stuart, J. S. & Binzel, R. P. Bias-corrected population, size distribution, and impact hazard for the near-Earth objects. *Icarus* **170**, 295–311 (2004).
- Johnson, B. C. & Melosh, H. J. Formation of spherules in impact produced vapor plumes. *Icarus* **217**, 416–430 (2012).
- Raizer, Y. P. Condensation of a cloud of vaporized matter expanding in vacuum. *Sov. Phys. JETP* **37**, 1229–1235 (1960).
- Melosh, H. J. & Vickery, A. M. Melt droplet formation in energetic impact events. *Nature* **350**, 494–497 (1991).
- O'Keefe, J. D. & Ahrens, T. J. The interaction of the Cretaceous/Tertiary extinction bolide with the atmosphere, ocean, and solid Earth. *Geol. Soc. Am. Spec. Pap.* **190**, 103–120 (1982).
- Smit, J. The global stratigraphy of the Cretaceous-Tertiary boundary impact ejecta. *Annu. Rev. Earth Planet. Sci.* **27**, 75–113 (1999).
- Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science* **208**, 1095–1108 (1980).
- Collins, G. S., Melosh, H. J., Morgan, J. V. & Warner, M. R. Hydrocode simulations of Chicxulub crater collapse and peak-ring formation. *Icarus* **157**, 24–33 (2002).
- Kyte, F. T., Shukolyukov, A., Lugmair, G. W., Lowe, D. R. & Byerly, G. R. Early Archean spherule beds: chromium isotopes confirm origin through multiple impacts of projectiles of carbonaceous chondrite type. *Geology* **31**, 283–286 (2003).
- Lowe, D. R. *et al.* Spherule beds 3.47–3.24 billion years old in the Barberton Greenstone Belt, South Africa: a record of large meteorite impacts and their influence on early crustal and biological evolution. *Astrobiology* **3**, 7–48 (2003).
- Yancey, T. E. & Guillemette, R. N. Carbonate accretionary lapilli in distal deposits of the Chicxulub impact event. *Geol. Soc. Am. Bull.* **120**, 1105–1118 (2008).
- Minton, D. A. & Malhotra, R. Dynamical erosion of the asteroid belt and implication for large impacts in the inner Solar System. *Icarus* **207**, 744–757 (2010).
- Lowe, D. R. & Byerly, G. R. Did LHB end not with a bang but a whimper? The geologic evidence. *41st Lunar Planet. Sci. Conf.* 2563 (2010).
- Gomes, R., Levinson, H. F., Tsiganis, K. & Morbidelli, A. Origin of the cataclysmic Late Heavy Bombardment period of the terrestrial planets. *Nature* **435**, 466–469 (2005).
- Botke, W. F. *et al.* An Archaean heavy bombardment from a destabilized extension of the asteroid belt. *Nature* <http://dx.doi.org/10.1038/nature10967> (this issue).
- Rasmussen, B. & Koeberl, C. Iridium anomalies and shocked quartz in a Late Archean spherule layer from Pilbara craton: new evidence for a major asteroid impact at 2.63 Ga. *Geology* **32**, 1029–1032 (2004).
- Kohl, I., Simonson, B. M. & Berke, M. Diagenetic alteration of impact spherules in the Neoproterozoic Monteville layer, South Africa. *Geol. Soc. Am. Spec. Pap.* **405**, 57–73 (2006).
- Goderis, S. *et al.* Geochemical correlation of two late Archean impact spherule layers between South Africa and Western Australia: the Paraburdoo-Reivilo link. *43rd Lunar Planet. Sci. Conf.* 1882 (2011).
- Hassler, S. W., Simonson, B. M., Sumner, D. Y. & Bodin, L. Paraburdoo spherule layer (Hamersley Basin, Western Australia): distal ejecta from the fourth large impact near the Archean-Proterozoic boundary. *Geology* **39**, 307–310 (2011).
- Simonson, B. M. *et al.* Geochemistry of 2.63–2.49 Ga impact spherule layers and implications for stratigraphic correlations and impact processes. *Precamb. Res.* **175**, 51–76 (2009).
- Simonson, B. M. Geological evidence for a strewn field of impact spherules in the early Precambrian Hamersley basin of Western Australia. *Geol. Soc. Am. Bull.* **104**, 829–839 (1992).
- Glikson, A. & Allen, C. Iridium anomalies and fractionated siderophile element patterns in impact ejecta, Brockman Iron Formation, Hamersley basin, Western Australia: evidence for a major asteroid impact in simatic crustal regions of the early Proterozoic Earth. *Earth Planet. Sci. Lett.* **220**, 247–264 (2004).
- Simonson, B. M., Sumner, D. Y., Beukes, N. J., Johnson, S. & Gutzmer, J. Correlating multiple Neoproterozoic-Paleoproterozoic impact spherule layers between South Africa and Western Australia. *Precamb. Res.* **169**, 100–111 (2009).
- Chadwick, B., Claeys, P. & Simonson, B. New evidence for a large Palaeoproterozoic impact: spherules in a dolomite layer in the Ketilidian orogeny, South Greenland. *J. Geol. Soc. Lond.* **158**, 331–340 (2001).
- Shukolyukov, A. & Lugmair, G. W. Isotopic evidence for the Cretaceous-Tertiary impactor and its type. *Science* **282**, 927–930 (1998).
- Glass, B. P., Koeberl, C., Blum, J. D. & McHugh, M. G. Upper Eocene tektite and impact ejecta layer on the continental slope off New Jersey. *Meteorit. Planet. Sci.* **33**, 229–241 (1998).
- Kyte, F. T., Shukolyukov, A., Hildebrand, A. R., Lugmair, G. W. & Hanova, J. Chromium isotopes in Late Eocene impact spherules indicate a likely asteroid belt provenance. *Earth Planet. Sci. Lett.* **302**, 279–286 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank B. Simonson and C. Chapman, whose comments improved this work. We acknowledge support from NASA.

**Author Contributions** B.C.J. developed the methods used in this work with the guidance of H.J.M. Both authors contributed to the conclusions presented in this work.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to B.C.J. ([johns477@purdue.edu](mailto:johns477@purdue.edu)).

# An Archaean heavy bombardment from a destabilized extension of the asteroid belt

William F. Bottke<sup>1</sup>, David Vokrouhlický<sup>1,2</sup>, David Minton<sup>1,3</sup>, David Nesvorný<sup>1</sup>, Alessandro Morbidelli<sup>1,4</sup>, Ramon Brasser<sup>1,4,5</sup>, Bruce Simonson<sup>6</sup> & Harold F. Levison<sup>1</sup>

The barrage of comets and asteroids that produced many young lunar basins (craters over 300 kilometres in diameter) has frequently been called the Late Heavy Bombardment<sup>1</sup> (LHB). Many assume the LHB ended about 3.7 to 3.8 billion years (Gyr) ago with the formation of Orientale basin<sup>2,3</sup>. Evidence for LHB-sized blasts on Earth, however, extend into the Archaean and early Proterozoic eons, in the form of impact spherule beds: globally distributed ejecta layers created by Chicxulub-sized or larger cratering events<sup>4</sup>. At least seven spherule beds have been found that formed between 3.23 and 3.47 Gyr ago, four between 2.49 and 2.63 Gyr ago, and one between 1.7 and 2.1 Gyr ago<sup>5–9</sup>. Here we report that the LHB lasted much longer than previously thought, with most late impactors coming from the E belt, an extended and now largely extinct portion of the asteroid belt between 1.7 and 2.1 astronomical units from Earth. This region was destabilized by late giant planet migration<sup>10–13</sup>. E-belt survivors now make up the high-inclination Hungaria asteroids<sup>14,15</sup>. Scaling from the observed Hungaria asteroids, we find that E-belt projectiles made about ten lunar basins between 3.7 and 4.1 Gyr ago. They also produced about 15 terrestrial basins between 2.5 and 3.7 Gyr ago, as well as around 70 and four Chicxulub-sized or larger craters on the Earth and Moon, respectively, between 1.7 and 3.7 Gyr ago. These rates reproduce impact spherule bed and lunar crater constraints.

When a large impactor strikes the Earth, it produces a vapour-rich ejecta plume containing numerous sand-sized melt droplets, most of which rise above the atmosphere. Eventually the droplets cool and fall back, forming a global layer that can be several millimetres to many centimetres thick for roughly Chicxulub-sized or larger impact events<sup>4</sup>. These layers have been identified in particular Archaean and early Proterozoic terrains that have been extensively searched, although preservation biases and incomplete sampling are still possibilities. The characteristics of the layers and the spherules themselves suggest that they are distal rather than proximal ejecta. They potentially provide us with a comprehensive record of large ancient impact events, even if their source craters were eliminated long ago.

Although the precise projectile size needed to form a global spherule bed is unknown, all Archaean and early Proterozoic beds are as thick as or thicker than those associated with the 65-million-year-old, 180-km-diameter Chicxulub crater. In comparison, the 35-million-year-old, 100-km-diameter Popigai crater, perhaps the second-largest crater known from the Phanerozoic, formed a distal spherule bed that is less than 0.1 mm thick<sup>6,9</sup>. Spherule beds as thin as that have yet to be detected on ancient terrains.

The known ancient beds argue for an intense, protracted phase of late terrestrial bombardment<sup>5–9</sup>. Curiously, these enormous blasts have no obvious source, even though many occurred relatively soon after the formation of the 930-km-diameter lunar basin Orientale (see ref. 13, for example). This makes us suspect that a key aspect of the LHB has been missed.

The best-developed dynamical model of the LHB, referred to here as the Nice model<sup>11,10</sup>, suggests that late giant planet migration drove resonances inward across the primordial main asteroid belt region. This event not only pushed numerous asteroids onto planet-crossing orbits, but also set up the current resonance structure of the main asteroid belt<sup>11,12</sup>. We use this framework to explore a possible lost source of late-LHB impactors.

The main asteroid belt's inner boundary is currently set by the  $\nu_6$  secular resonance at 2.1 AU (one astronomical unit is approximately the Earth–Sun distance); objects entering this resonance have their eccentricities pumped up to planet-crossing values in less than a million years<sup>16</sup>. Before the LHB, the giant planets and their associated secular resonances were in different locations, with the only remaining natural inner boundary being the Mars-crossing zone. Accordingly, the main asteroid belt may have once stretched into the E-belt zone as far as 1.7 AU.

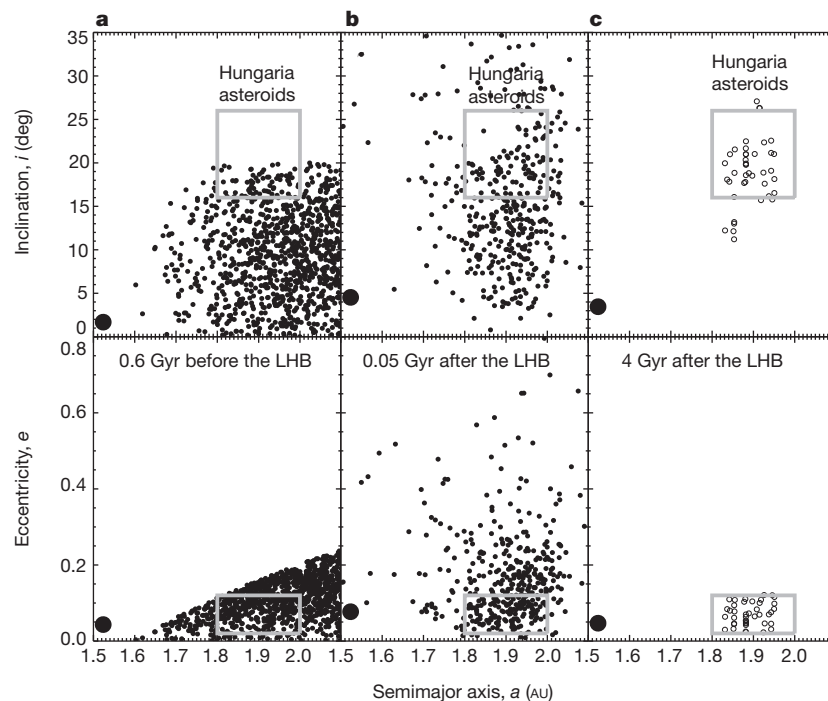
To determine what would have happened to E-belt objects before and after planet migration, we tracked four sets of 1,000 model asteroids, with the population started with semimajor axes between 1.7–2.1 AU and main-asteroid-belt-like eccentricities and inclinations (Fig. 1a). In the pre-LHB phase, we assumed Nice-model-like initial conditions for the planets: Venus and Earth were on their current orbits, while the giant planets were on circular and nearly coplanar orbits between 5.4 and 11.7 AU (ref. 12). Our primary variable was the initial eccentricity of Mars, which conceivably could have been different at this time<sup>17</sup>. We set its maximum osculating value to  $e_{\text{max}}^{\text{Mars}} = 0.025, 0.05, 0.12$  (its current value) and 0.17. We found most of our test asteroids were stable for 0.6 Gyr; losses were less than 15% for all but the  $e_{\text{max}}^{\text{Mars}} = 0.17$  run (Figs 1 and 2). The bodies that did escape generally came from the periphery of the Mars-crossing zone, where they were perturbed onto planet-crossing orbits via interactions with Mars.

Next, in the LHB phase, the giant planets and their resonances were assumed to migrate to their current orbits in less than a million years. This behaviour was approximated in our model by instantaneously ‘jumping’ the giant planets to their current orbits. Numerical models show that resonances must have swept rapidly across the primordial main asteroid belt from the outside in, depleting it by about 75% (refs 11–13). They may also have allowed Mars to achieve its current orbital eccentricity via secular resonant coupling between the terrestrial and giant planets<sup>17</sup>.

The sudden appearance of the  $\nu_6$  secular resonance at its current location, along with related resonances in the same region, destabilized E-belt asteroids by exciting their eccentricities and inclinations (Fig. 1b). Over the next 4 Gyr, these effects drove nearly all E-belt asteroids onto planet-crossing orbits (Fig. 2). En route, many passed through or near the Hungaria asteroid region, located at high inclinations between 1.8–2.0 AU (Fig. 1b)<sup>14,15</sup>.

<sup>1</sup>Southwest Research Institute and NASA Lunar Science Institute, 1050 Walnut Street, Suite 300, Boulder, Colorado 80302, USA. <sup>2</sup>Institute of Astronomy, Charles University, V. Holešovičkách 2, 18000 Prague 8, Czech Republic. <sup>3</sup>Department of Earth and Atmospheric Sciences, Purdue University, 550 Stadium Mall Drive, West Lafayette, Indiana 47907, USA. <sup>4</sup>Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, Laboratoire Cassiopée, BP 4229, 06304 Nice Cedex 4, France. <sup>5</sup>Institute of Astronomy and Astrophysics, Academia Sinica, PO Box 23 141, Taipei 106, Taiwan. <sup>6</sup>Geology Department, Oberlin College, Oberlin, Ohio 44074, USA.





**Figure 1 | Snapshots of the evolution of the E-belt population over time.** **a**, 0.6 Gyr before the LHB. Here the filled circles show 1,000 randomly created test asteroids. They were selected according to a uniform distribution of semimajor axes  $a$  at 1.6–2.1 AU and a main-asteroid-belt-like Gaussian distribution in eccentricity  $e$  and inclination  $i$ , the latter having peaks at  $0.15^\circ$  and  $8.5^\circ$ , respectively, and standard deviations of  $0.07^\circ$  and  $7^\circ$ , respectively<sup>13</sup>. All test bodies initially placed on Mars-crossing orbits were rejected, and almost no objects achieved  $a < 1.7$  AU. This yielded a population that was equivalent to 16% of the primordial main-asteroid-belt population between 2.1 and 3.25 AU. The planets were started on their pre-LHB orbits as defined in the main text, with  $e_{\text{Mars}}^{\text{max}} = 0.05$ . All bodies were tracked using the symplectic integration code SWIFT-RMVS3<sup>29</sup>.

The Hungaria population is the quasi-stable reservoir of small asteroids closest to the terrestrial planet region. Bracketed by multiple resonances, the region is dynamically ‘sticky’—objects finding a way in often take a long time to come back out. This is reflected in our model runs, with our non-planet-crossing survivors after 4 Gyr always found on Hungaria-like orbits (Fig. 1c). We infer from this that if the initial E belt were large enough, it could have produced the Hungaria asteroids.

The current Hungaria population is comprised of a single E-type asteroid family, with E-types thought to have enstatite chondrite-like surfaces, set among a diverse background asteroid population (for example, E-, X-, S- and C-type asteroids)<sup>14,15</sup>. By accounting for their expected lunar impact velocities, which can be quite high (20% hit at over  $30 \text{ km s}^{-1}$ ), we estimate that the Hungaria asteroids have  $4 \pm 2$  objects capable of forming basins. This is tiny compared to the 7,500 or so that exist in the main asteroid belt<sup>18</sup> (see Supplementary Information).

Approximately 0.1–0.4% of our E-belt asteroids found refuge in the Hungaria region (Fig. 2). These values are probably upper limits; additional objects may have been eliminated by collisional evolution or migration onto planet-crossing orbits through a combination of Yarkovsky thermal forces and resonances<sup>19</sup>. We apply a depletion factor of 1.5 to account for these effects. Scaling from the known Hungaria asteroids, we estimate that the E belt’s population just before the LHB was approximately 0.2–0.8 times that of the current main-asteroid-belt population<sup>18</sup>. Interestingly, the larger values yield a population density consistent with the primordial main asteroid belt just before the LHB<sup>13</sup>. This suggests that the E belt does not need to be exceptional to reproduce the Hungaria asteroids.

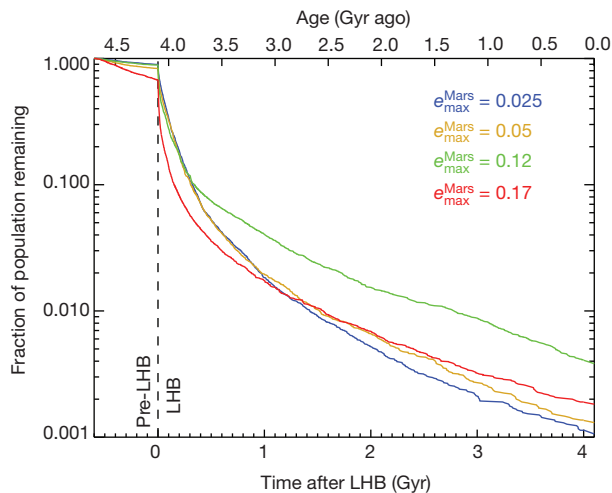
We find that a combination of our  $e_{\text{Mars}}^{\text{max}} = 0.025$  and 0.05 runs produce, on average, nine or ten lunar basins during the LHB (see

The large black circle is Mars. **b**, 0.05 Gyr after the start of the LHB. The planets are now on their current orbits. The population has dropped from 83% just before the LHB to 44% now. Many bodies have been driven onto orbits similar to those of the Hungaria asteroids (grey box;  $1.78 < a < 2.0$  AU;  $e < 0.15$ , and  $16^\circ < i < 26^\circ$ ; refs 14 and 15). **c**, 4 Gyr after the start of the LHB. The open circles are 41 clones of the original test asteroids and represent 0.14% of the original population. Our clones were created by adding a uniformly distributed random number between  $-1 \times 10^{-6}$  AU and  $+1 \times 10^{-6}$  AU to each of the position coordinates of our test bodies at a set time. Here the surviving test bodies were cloned ten times when 90% of the initial population was lost and three times once 90% of those clones were lost. Most reside in the Hungaria region.

Supplementary Information). The main asteroid belt’s contribution is about three, so together we get 12 or 13 lunar basins. The E belt dominates, despite its small size, because its asteroids have a probability of hitting the Earth and Moon that is ten times higher than those originating in the main belt<sup>13</sup>. In an end-member model, where E-belt and main-belt asteroids produce all of the lunar LHB, this would place its start near the formation time of the Nectaris basin<sup>2,3</sup>. As a check, we compared crater counts on Nectaris terrains to our expected crater populations and found an excellent match (see Supplementary Information). This could imply that comets are a minor player in the bombardment of the Earth and Moon during the LHB, as suggested by certain lines of evidence (such as the shape of the lunar crater size frequency distributions<sup>20</sup> and the inferred nature of basin projectiles<sup>21</sup>; see Supplementary Information).

Our LHB-era lunar basins form over an approximately 400-million-year interval, much longer than previous estimates<sup>1</sup> (Fig. 3). Thus, if the Orientale basin formed 3.7–3.8 Gyr ago, the LHB starting time should be 4.1–4.2 Gyr ago. These ages are intriguing because many lunar samples were modified by ancient impact heating events between 3.7 and 4.1 Gyr ago<sup>22,23</sup>. Similarly, in the asteroid belt,  $^{39}\text{Ar}$ – $^{40}\text{Ar}$  shock degassing ages for eucrite and H-chondrite meteorites show a paucity of ages 4.1–4.4 Gyr ago and numerous ages 3.5–4.1 Gyr ago<sup>23,24</sup>. For Mars, the data are meagre, but it is interesting to note that the crystallization age of Martian meteorite ALH84001 is about 4.1 Gyr ago<sup>25</sup>. The Martian shergottite source region also appears to have been disturbed at the same time<sup>26</sup>.

Lunar basins and craters formed before 4.1–4.2 Gyr ago would presumably come from leftover planetesimals in the terrestrial planet region<sup>27</sup> and pre-LHB refugees from the primordial E belt and main

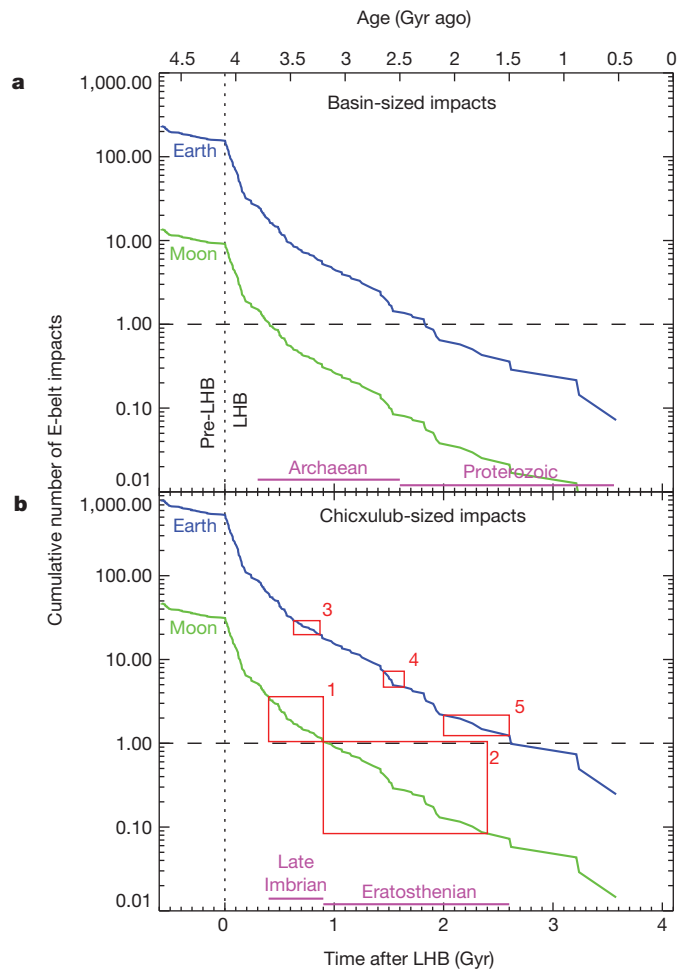


**Figure 2 | Decay curves for our E-belt runs before and after the LHB.** The blue, gold, green and red lines correspond to the different maximum eccentricities of Mars before the LHB, with  $e_{\text{Mars}}^{\text{max}} = 0.025, 0.05, 0.12$  (current value) and  $0.17$ , respectively. The test asteroids corresponded to starting E-belt populations with 18%, 16%, 9.3% and 5.4% of the primordial main-asteroid-belt population, respectively. Many asteroids eliminated before the LHB had orbits near the Mars-crossing region; they were scattered onto planet-crossing orbits via Mars perturbations. The  $e_{\text{Mars}}^{\text{max}} = 0.12$  run was found to decay more slowly than the other runs because a larger fraction of its test asteroids were dynamically pushed near or through the quasi-stable Hungaria region. The effects of collisional evolution<sup>18</sup> and non-gravitational (Yarkovsky) forces on the objects<sup>19</sup> were not included, although they must affect our results. The former is most likely to affect the pre-LHB phase when the initial populations were massive. Tests for the latter indicate the amount of additional depletion produced for objects over 10 km in diameter in the LHB era is less than a factor of two, although it is more substantial for smaller asteroids. Taken together, we estimate that an extra depletion factor of 1.5 should reasonably account for these effects, though we consider this value to be conservative. In addition, although these decay rates do not suffer from small-number statistics, the real population does (that is, only  $4 \pm 2$  basin-forming projectiles now exist in the current Hungaria population). Using a Monte Carlo code to track asteroid depletion, we find that our estimates of the initial E-belt population could easily vary by an additional factor of two (see Supplementary Information). If these factors worked in the right direction, it could allow our more eccentric Mars cases to also match LHB constraints.

asteroid belt (Fig. 3). They would have hit the Moon while the giant planets resided on nearly circular, largely coplanar orbits, with inner Solar System resonances too weak to produce the same degree of dynamical excitation as observed today among near-Earth asteroids. This result is reflected in our model lunar impact velocities, whose median values double from  $9 \text{ km s}^{-1}$  to  $21 \text{ km s}^{-1}$  once the LHB begins (see Supplementary Information).

Evidence for such a velocity change may exist on the Moon. Lunar crater size distributions on Nectaris terrains are found to have the same basic shape as those on the most ancient lunar terrains with diameters between 20 km and 150 km, but Nectaris craters are larger by 30–40% (ref. 28). This shift is consistent with impact velocities in the LHB-era increasing by a factor of about two (see Supplementary Information).

Our model results also demonstrate the need for an LHB. The E-belt model runs shown in Fig. 2 have steep decay rates as the LHB begins, but they then transition to shallow ones for the last billion years. This implies that shifting the start of the LHB to early Solar System times, say 4.5 Gyr ago, would only increase the size of the primordial E-belt population by about 30%, because it still has to match constraints from the current Hungaria population, but it would take away the E belt's ability to produce late lunar basins. Accordingly, if the giant planets reach their current orbital configuration much earlier than 4.2 Gyr ago, the E belt and main asteroid belt cannot form lunar basins like Imbrium and Orientale at 3.7–3.9 Gyr ago.



**Figure 3 | The E-belt impactor flux on the Earth and Moon.** **a**, The number of E-belt impacts making basin-sized craters (diameters over 300 km) on Earth and the Moon. The curves represent the combined results of the  $e_{\text{Mars}}^{\text{max}} = 0.025$  and  $0.05$  runs. They were normalized assuming that about nine lunar basins form during the LHB and that the age of the last lunar basin (Orientale) is 3.7 Gyr ago<sup>3</sup>. **b**, The number of E-belt impacts making Chicxulub-sized craters (diameters over 160 km) on Earth and the Moon. The curves were scaled up by a factor of 3.4, approximately the ratio of basin- and Chicxulub-forming projectiles in the main-asteroid-belt size distribution<sup>18,30</sup>; see also <http://www.lpl.arizona.edu/tekton/crater.html>. The red boxes denote time intervals with constraints. For the lunar Late-Imbrian era (box 1; 3.2–3.7 Gyr ago), there are three such craters observed (Iridium, Humboldt, Tsiolkovskiy, with diameters of 260, 207 and 180 km, respectively), while for the Eratosthenian era (box 2; 1.5–3.2 Gyr ago), there is one observed (Hausen, 167 km in diameter)<sup>23</sup>. Hausen might also be a Late Imbrian-era crater<sup>31</sup>. The remainder correspond to terrestrial impact spherule beds from specific Archaean and early Proterozoic terrains that have been extensively searched: at least seven beds between 3.23 and 3.47 Gyr ago (box 3), four beds between 2.49 and 2.63 Gyr ago (box 4), and one bed between 1.7 and 2.1 Gyr ago (box 5)<sup>5–9</sup>. The Chicxulub-sized craters Vredefort (2.02 Gyr ago) and Sudbury (1.85 Gyr ago) formed in box 5, so the true value should be 2 (ref. 3). No spherule beds have been found between 0.6 and 1.7 Gyr ago, but this time interval has not been extensively explored for impact spherules<sup>6</sup>. Over the same time intervals, our model results are essentially identical; for boxes 1 to 5, we obtain  $3 \pm 2, 1 \pm 1, 9 \pm 3, 3 \pm 2$  and  $1 \pm 1$ , respectively.

The E belt continued to produce large lunar impacts well after the conventional end of the LHB at 3.7–3.8 Gyr ago, with three roughly Chicxulub-sized or larger craters made on Late-Imbrian-era terrains (3.2–3.7 Gyr ago) and perhaps one on Eratosthenian-era terrains (1.5–3.2 Gyr ago); see ref. 2 (<http://ser.sese.asu.edu/GHM/>) and ref. 3 (Fig. 3). These values match observations. In contrast, very few comparable-sized impacts are expected to come from ejected main-belt asteroids or comets over these times<sup>13</sup>.

The ratio of the gravitational cross-sections of Earth and the Moon found using our E-belt encounter velocities is about 17:1. Thus, as predicted by our simulations, the existence of four Chicxulub-sized lunar craters younger than 3.7–3.8 Gyr implies that  $68 \pm 8$  similar impacts should have taken place on Earth over a comparable formation period. These impact rates yield  $9 \pm 3$ ,  $3 \pm 2$ ,  $1 \pm 1$  and  $1 \pm 1$  events over 3.23–3.47, 2.49–2.63, 1.7–2.1 and 0.6–1.7 Gyr ago, respectively, enough to reproduce the known Archaean and early Proterozoic spherule bed data (Fig. 3b).

The largest Archaean-era blasts rivalled those that formed lunar basins during the Nectarian- and Early Imbrian eras. We calculate that  $15 \pm 4$  basin-forming impactors struck Earth between 2.5 and 3.7 Gyr ago (Fig. 3a). Some may even have been as big as the one that formed the Orientale basin<sup>4,5</sup>. The terrestrial consequences of these mammoth Archaean events have yet to be explored, but we suspect that they may have affected the evolution of life and our biosphere in profound ways.

Received 29 November 2011; accepted 17 February 2012.

Published online 25 April 2012.

- Gomes, R., Levison, H. F., Tsiganis, K. & Morbidelli, A. Origin of the cataclysmic Late Heavy Bombardment period of the terrestrial planets. *Nature* **435**, 466–469 (2005).
- Wilhelms, D. E. The geologic history of the Moon. *US Geol. Surv. Prof. Pap.* **1348**, 1987 (1987).
- Stöffler, D. & Ryder, G. Stratigraphy and isotope ages of lunar geologic units: chronological standard for the inner Solar System. *Space Sci. Rev.* **96**, 9–54 (2001).
- Johnson, B. C. & Melosh, H. J. Impact spherules as a record of ancient heavy bombardment of Earth. *Nature* <http://dx.doi.org/10.1038/nature10982> (this issue).
- Lowe, D. R. *et al.* Spherule beds 3.47–3.24 billion years old in the Barberton Greenstone belt, South Africa: a record of large meteorite impacts and their influence on early crustal and biological evolution. *Astrobiology* **3**, 7–48 (2003).
- Simonson, B. M. & Glass, B. P. Spherule layers—records of ancient impacts. *Annu. Rev. Earth Planet. Sci.* **32**, 329–361 (2004).
- Lowe, D. R. & Byerly, G. R. Did LHB end not with a bang but a whimper? The geologic evidence. *Lunar Planet. Sci.* **41**, 2563 (2010).
- Hassler, S. W., Simonson, B. M., Sumner, D. Y. & Bodin, L. Paraburdoo spherule layer (Hamersley Basin, Western Australia): distal ejecta from a fourth large impact near the Archaean-Proterozoic boundary. *Geology* **39**, 307–310 (2011).
- Glass, B. P. & Simonson, B. M. Distal impact ejecta layers: spherules and more. *Elements* **8**, 43–48 (2012).
- Tsiganis, K., Gomes, R., Morbidelli, A. & Levison, H. F. Origin of the orbital architecture of the giant planets of the Solar System. *Nature* **435**, 459–461 (2005).
- Minton, D. A. & Malhotra, R. Secular resonance sweeping of the main asteroid belt during planet migration. *Astrophys. J.* **732**, 53 (2011).
- Morbidelli, A., Brasser, R., Gomes, R., Levison, H. F. & Tsiganis, K. Evidence from the asteroid belt for a violent past evolution of Jupiter's orbit. *Astron. J.* **140**, 1391–1401 (2010).
- Minton, D. A. & Malhotra, R. Dynamical erosion of the asteroid belt and implications for large impacts in the inner Solar System. *Icarus* **207**, 744–757 (2010).
- Warner, B. D., Harris, A. W., Vokrouhlický, D., Nesvorný, D. & Bottke, W. F. Analysis of the Hungaria asteroid population. *Icarus* **204**, 172–182 (2009).
- Milani, A., Knežević, Z., Novaković, B. & Cellino, A. Dynamics of the Hungaria asteroids. *Icarus* **207**, 769–794 (2010).
- Bottke, W. F. *et al.* Debaised orbital and absolute magnitude distribution of the near-Earth objects. *Icarus* **156**, 399–433 (2002).
- Brasser, R., Morbidelli, A., Gomes, R., Tsiganis, K. & Levison, H. F. Constructing the secular architecture of the Solar System. II: The terrestrial planets. *Astron. Astrophys.* **507**, 1053–1065 (2009).
- Bottke, W. F. *et al.* Linking the collisional history of the main asteroid belt to its dynamical excitation and depletion. *Icarus* **179**, 63–94 (2005).
- Bottke, W. F., Vokrouhlický, D., Rubincam, D. P. & Nesvorný, D. The Yarkovsky and YORP effects: implications for asteroid dynamics. *Annu. Rev. Earth Planet. Sci.* **34**, 157–191 (2006).
- Strom, R. G., Malhotra, R., Ito, T., Yoshida, F. & Kring, D. A. The origin of planetary impactors in the inner Solar System. *Science* **309**, 1847–1850 (2005).
- Kring, D. A. & Cohen, B. A. Cataclysmic bombardment throughout the inner solar system 3.9–4.0 Ga. *J. Geophys. Res.* **107** (E2) 5009 (2002).
- Norman, M. D., Duncan, R. A. & Huard, J. J. Imbrium provenance for the Apollo 16 Descartes terrain: argon ages and geochemistry of lunar breccias 67016 and 67455. *Geochim. Cosmochim. Acta* **74**, 763–783 (2010).
- Bogard, D. D. Impact ages of meteorites: a synthesis. *Meteoritics* **30**, 244–268 (1995).
- Bogard, D. D. K-Ar ages of meteorites: clues to parent-body thermal histories. *Chem. Erde Geochem.* **71**, 207–226 (2011).
- Lapen, T. J. *et al.* A younger age for ALH84001 and its geochemical link to shergottite sources in Mars. *Science* **328**, 347–351 (2010).
- Bogard, D. D. & Park, J. <sup>39</sup>Ar–<sup>40</sup>Ar dating of the Zagami Martian shergottite and implications for magma origin of excess <sup>40</sup>Ar. *Meteorit. Planet. Sci.* **43**, 1113–1126 (2008).
- Bottke, W. F., Levison, H. F., Nesvorný, D. & Dones, L. Can planetesimals left over from terrestrial planet formation produce the lunar Late Heavy Bombardment? *Icarus* **190**, 203–223 (2007).
- Marchi, S., Bottke, W. B., Kring, D. A. & Morbidelli, A. The onset of the lunar cataclysm as recorded in its ancient crater populations. *Earth Planet. Sci. Lett.* **325–326**, 27–38 (2012).
- Levison, H. F. & Duncan, M. J. The long-term dynamical behavior of short-period comets. *Icarus* **108**, 18–36 (1994).
- Melosh, H. J. *Impact Cratering: A Geologic Process* Ch. 7 (Oxford University Press, 1989).
- McEwen, A. S. *et al.* Galileo observations of post-Imbrium lunar craters during the first Earth-Moon flyby. *J. Geophys. Res.* **98**, 17,207–17,231 (1993).

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank B. Cohen, C. Chapman, M. Čuk, L. Dones, D. Kring, S. Marchi, R. Malhotra, M. Norman, E. Scott, J. Taylor and K. Walsh for discussions and comments. We also thank the University of Hawaii for sponsoring W.F.B.'s recent visit, during which these ideas were first developed. This project was supported by NASA's Lunar Science Institute (Center for Lunar Origin and Evolution, grant number NNA09DB32A). D.V.'s contribution was supported by the Grant Agency of the Czech Republic. A.M. and R.B. thank Germany's Helmholtz Alliance for providing support through their "Planetary Evolution and Life" programme. B.S.'s contribution was supported by NASA grant NNX08AI29G. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center.

**Author Contributions** W.F.B. developed the scenario of the E belt. The numerical runs were constructed by W.F.B. and D.V., with input from D.M., D.N. and H.F.L. Information on the initial conditions of the planets and the nature of late giant planet migration within the Nice model was provided by A.M., R.B., H.F.L. and D.N. Information on the nature of sweeping resonances, how they affected the main belt, and what happened afterwards was provided by D.M., A.M. and R.B. Additional numerical runs not shown here were performed by A.M. and R.B. Information on the nature of impact spherules and their context was provided by B.S. All authors participated in numerous discussions.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to W.F.B. ([bottke@boulder.swri.edu](mailto:bottke@boulder.swri.edu)).



# Spin–orbital separation in the quasi–one–dimensional Mott insulator $\text{Sr}_2\text{CuO}_3$

J. Schlappa<sup>1,2</sup>, K. Wohlfeld<sup>3</sup>, K. J. Zhou<sup>1†</sup>, M. Mourigal<sup>4</sup>, M. W. Haverkort<sup>5</sup>, V. N. Strocov<sup>1</sup>, L. Hozoi<sup>3</sup>, C. Monney<sup>1</sup>, S. Nishimoto<sup>3</sup>, S. Singh<sup>6†</sup>, A. Revcolevschi<sup>6</sup>, J.-S. Caux<sup>7</sup>, L. Patthey<sup>1,8</sup>, H. M. Rønnow<sup>4</sup>, J. van den Brink<sup>3</sup> & T. Schmitt<sup>1</sup>

**When viewed as an elementary particle, the electron has spin and charge. When binding to the atomic nucleus, it also acquires an angular momentum quantum number corresponding to the quantized atomic orbital it occupies. Even if electrons in solids form bands and delocalize from the nuclei, in Mott insulators they retain their three fundamental quantum numbers: spin, charge and orbital<sup>1</sup>. The hallmark of one-dimensional physics is a breaking up of the elementary electron into its separate degrees of freedom<sup>2</sup>. The separation of the electron into independent quasi-particles that carry either spin (spinons) or charge (holons) was first observed fifteen years ago<sup>3</sup>. Here we report observation of the separation of the orbital degree of freedom (orbiton) using resonant inelastic X-ray scattering on the one-dimensional Mott insulator  $\text{Sr}_2\text{CuO}_3$ . We resolve an orbiton separating itself from spinons and propagating through the lattice as a distinct quasi-particle with a substantial dispersion in energy over momentum, of about 0.2 electronvolts, over nearly one Brillouin zone.**

It was pointed out in the 1970s that in a solid not only the charge and spin of electrons can become ordered—leading to magnetism—but also the electrons' orbital degree of freedom<sup>1</sup>. This observation sparked a field that has gone on to produce a number of important results. Although a physical electron combines spin, charge and orbital, theoretically an electron can be considered a bound state of the three independent, fundamental quasi-particles: a spinon, carrying the electron's spin; a holon (or chargon), carrying its charge; and an orbiton, carrying its orbital degree of freedom.

A remarkable and fundamental property of one-dimensional (1D) systems is that electronic excitations break up into deconfined spinons and holons. This was predicted decades ago (ref. 2 and references therein) and confirmed in the mid 1990s by angle-resolved photoemission spectroscopy experiments<sup>3–5</sup>. The spin–charge separation is an example of particle fractionalization, a phenomenon in which the quantum numbers of quasi-particles are not multiples of those of the elementary particle, but fractions. This effect is one of the most unusual manifestations of collective quantum physics of interacting particles and is a profound concept that has found its way into a number of theories, for example that describing high-temperature superconductivity in copper oxides<sup>6,7</sup>.

To search for the further fractionalization of the electron, we consider the excitation of a copper orbital degree of freedom in the antiferromagnetic spin-chain compound  $\text{Sr}_2\text{CuO}_3$ . The spin–orbital separation process that we are looking for is analogous to the spin–charge separation mechanism (Fig. 1b). The latter occurs, for instance, when an electron is annihilated, removing a single spin and leaving behind a hole in the antiferromagnetic chain. This hole can start to propagate freely only after exciting one spinon (a domain wall in the antiferromagnetic chain). Subsequently, the spinon can delocalize and

separate itself completely from the holon. When instead of creating a hole, as typically is done in a photoemission experiment, an electron is excited from one copper  $3d$  orbital to another, the phenomenon of spin–orbital separation can in principle occur (Fig. 1a). The orbiton created in this manner may also deconfine after exciting a spinon, thus splitting the electron into its orbital and spin degrees of freedom<sup>8</sup>.

Here we use high-resolution resonant inelastic X-ray scattering (RIXS) to search experimentally for spin–orbital separation in the quasi-1D copper oxide  $\text{Sr}_2\text{CuO}_3$  (for material details, see Supplementary Information, section 1). We observe deconfinement of the spinon and orbiton during orbital excitation from the ground-state copper  $3d\ x^2 - y^2$  orbital to an excited copper  $3d\ xy$  or  $xz$  orbital (Fig. 1c–e). For simplicity, we will from henceforth use the so-called 'hole' language: although nominally there are nine electrons in the  $3d$  orbitals of the  $\text{Cu}^{2+}$  ion in  $\text{Sr}_2\text{CuO}_3$ , by using the 'electron–hole' transformation we can map the problem onto an effective system with one particle occupying a single  $3d$  orbital (Supplementary Information, section 2).

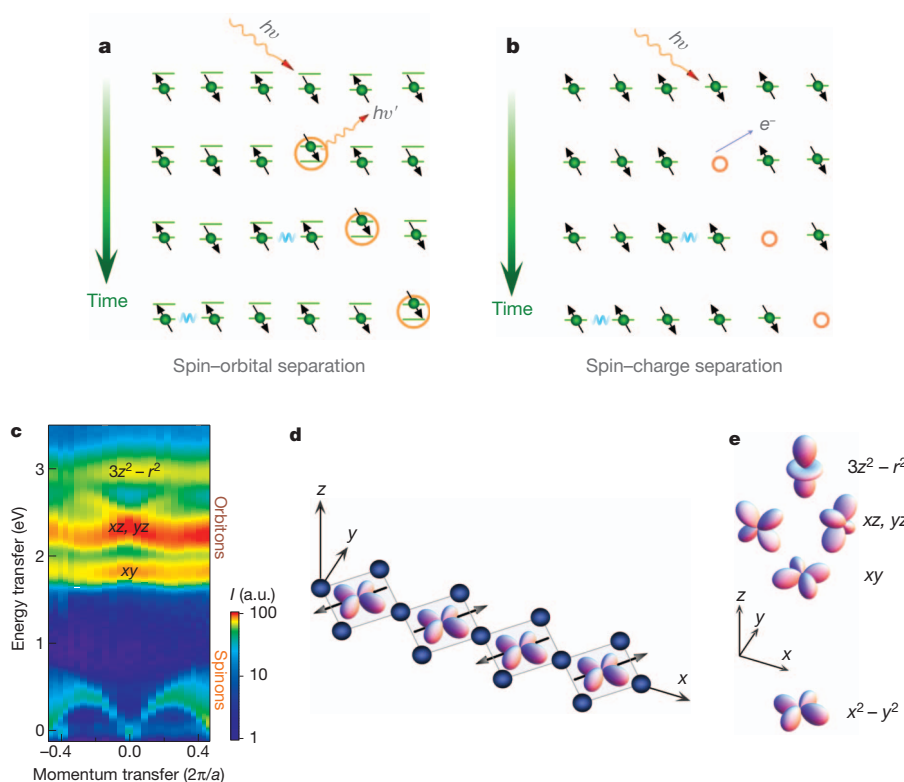
We measure an orbiton dispersion that is almost as large as the dispersion of the two-spinon continuum at low energies. As for spin–charge separation<sup>3–5</sup>, the orbiton dispersion has periodicity  $\pi$  (see Fig. 1c and see discussion below), which indicates the presence of an orbiton liberated from the spinon.

We measured the orbital excitations of  $\text{Sr}_2\text{CuO}_3$  using RIXS at the  $L_3$  edge of the copper ion. RIXS is a second-order scattering technique and can excite transitions between the copper  $3d$  states of different symmetry (orbital excitations), owing to the involvement of two subsequent electric dipole transitions<sup>9,10</sup> (Supplementary Information, section 3). With the unique capability of RIXS also to probe spin excitations<sup>11–13</sup> and to vary the photon momentum transfer, the dispersion of orbital and spin excitations can be mapped out across the first Brillouin zone<sup>11–16</sup>. The experiments were carried out at the ADDRESS beamline of the Swiss Light Source at the Paul Scherrer Institut<sup>17,18</sup>.

For fixed momentum transfer,  $\mathbf{q}$ , along the chains, peaks in the RIXS spectrum at constant energy transfer reveal the presence of charge-neutral elementary excitations and are visible in the RIXS intensity map of  $\text{Sr}_2\text{CuO}_3$  across the copper  $L_3$  edge in Fig. 2a. The spectrum for which the incident energy was precisely tuned to the resonance maximum of the absorption spectrum is shown in Fig. 2b. In both plots, the excitations of the spin, orbital and charge degrees of freedom are indicated. The momentum dependence and, in particular, the dispersion of the spin and orbital excitations (Fig. 1c and Supplementary Fig. 2a) are indicative of their collective nature.

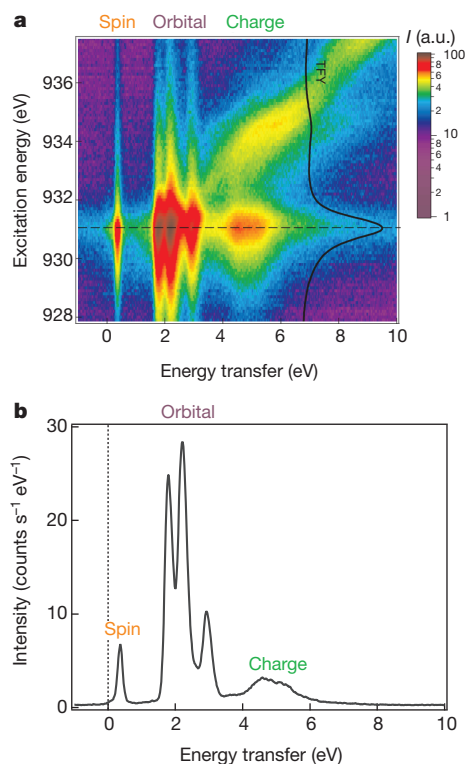
For energy transfers of up to  $\sim 0.8$  eV purely magnetic excitations are present, but the spectrum between  $\sim 1.5$  and  $\sim 3.5$  eV corresponds to excitations from the copper  $3d\ x^2 - y^2$  ground state to orbitals of  $xy$ ,  $xz/yz$  and  $3z^2 - r^2$  symmetry (Fig. 1d, e). These peaks correspond to

<sup>1</sup>Paul Scherrer Institut, Swiss Light Source, CH-5232 Villigen PSI, Switzerland. <sup>2</sup>Institut Methoden und Instrumentierung der Forschung mit Synchrotronstrahlung G-12, Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, D-12489 Berlin, Germany. <sup>3</sup>Institute for Theoretical Solid State Physics, IFW Dresden, Helmholtzstrasse 20, 01069 Dresden, Germany. <sup>4</sup>Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. <sup>5</sup>Max Planck Institute for Solid State Research, D-70569 Stuttgart, Germany. <sup>6</sup>CMMO - UMR 8182 - Bâtiment 410, Université Paris-Sud 11, 91405 Orsay Cedex, France. <sup>7</sup>Institute for Theoretical Physics, University of Amsterdam, Science Park 904, Postbus 94485, 1090 GL Amsterdam, The Netherlands. <sup>8</sup>SwissFEL, CH-5232 Villigen PSI, Switzerland. <sup>†</sup>Present addresses: Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK (K.J.Z.); Indian Institute of Science Education and Research, 900 NCL Innovation Park, Pashan, 411008 Pune, India (S.S.).



**Figure 1 | Spin–orbital separation process in an antiferromagnetic spin chain, emerging after exciting an orbital.** **a, b,** Sketches of spin–orbital separation (**a**) and spin–charge separation (**b**), generated in processes of RIXS and angle-resolved photoemission spectroscopy, respectively. The spin is represented by the arrow and the charge is represented by the green dot. In **a**, the lower and upper lines represent the ground-state and excited-state orbitals, respectively. **c**, RIXS intensity map of the dispersing spin and orbital excitations in  $\text{Sr}_2\text{CuO}_3$  as functions of photon momentum transfer along the

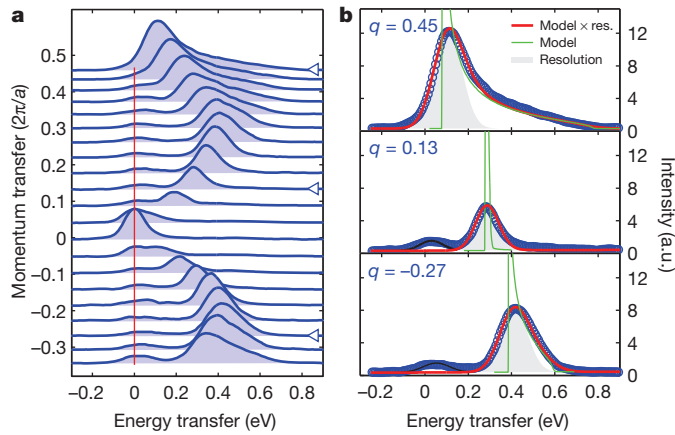
chains and photon energy transfer (for details, see main text and Supplementary Information, section 3). **a**,  $[100]$  lattice constant; a.u., arbitrary units. Data obtained at  $\Psi = 130^\circ$ . **d**, Geometry of the  $\text{CuO}_3$  chain, with the ground-state copper  $3d\ x^2 - y^2$  orbitals in the middle of each plaquette and oxygen sites at the plaquette corners. **e**, Orbital symmetries of  $x^2 - y^2$  and excited  $3d$  orbitals. In panels **c–e**, ‘hole’ language is used (Supplementary Information, section 2).



orbital excitations (called also  $d-d$  excitations), and not, for example, to charge transfer excitations, the intensity of which is non-zero for energies up to  $\sim 6$  eV but here for the case of  $L$ -edge RIXS is at least an order of magnitude lower<sup>10</sup>. The orbital assignment of these excitations was unambiguously verified by comparing their energy at  $\mathbf{q} = 0$  with *ab initio* quantum chemistry cluster calculations<sup>19</sup> (Supplementary Information, section 5, for detailed results).

Zooming into the ‘magnetic’ part of Fig. 1c, between 0 and 0.8 eV in energy transfer, reveals strongly dispersing spin excitations. At the lower boundary the dispersion has period  $\pi$ , and at the continuum (upper boundary) it has period  $2\pi$  (Fig. 3a). These RIXS data agree very well with recent inelastic neutron scattering studies on  $\text{Sr}_2\text{CuO}_3$  (ref. 20). The simultaneous presence in the spectrum of a lower edge with period  $\pi$  and an upper one with period  $2\pi$  indicates directly that in the spin chain the magnetic excitations with spin 1 break up and fractionalize into two-spinon (and higher-order) excitations that make up a continuum<sup>20</sup>, with each spinon having spin  $1/2$ . These spectra confirm that RIXS for magnetic excitations probes the well-known spin dynamical structure factor as theoretically predicted<sup>11–13</sup>, in agreement with recent studies on  $\text{TiOCl}$  (ref. 21). The excellent statistics of

**Figure 2 | Energy dependence of elementary excitations in  $\text{Sr}_2\text{CuO}_3$  observed with RIXS at the copper  $L_3$ -edge resonance.** **a**, RIXS map of  $\text{Sr}_2\text{CuO}_3$  (log scale) as a function of photon excitation energy (left axis) and energy transfer (bottom axis). The superimposed black curve shows the total fluorescence yield X-ray absorption spectrum. The dashed line marks the maximum energy of the copper  $L_3$ -edge resonance. **b**, RIXS line spectrum measured at the resonance maximum (along the dashed line in **a**). All data was obtained at a scattering angle of  $\Psi = 90^\circ$  and for momentum transfer along the chain of  $q = 0.189 \times 2\pi/a$  (Methods).



**Figure 3 | Dispersion of magnetic excitations: experimental data and simulation.** **a**, RIXS energy transfer spectra of the magnetic two-spinon continuum (see the spinon region in Fig. 1c). Data obtained at  $\Psi = 130^\circ$ . **b**, Fit (red line) of the experimental data (dots) with the two- and four-spinon dynamical structure factor<sup>22</sup> (green line), convolved with a Gaussian distribution (light-grey shaded region) to account for the total instrumental resolution of 140 meV. The low-energy peak at around 50 meV (black line) is elastic and phonon scattering. The values of momentum transfer for the selected spectra (shown in units of  $2\pi/a$ ) are indicated by arrows in **a**.

the data further allow for a direct comparison of the RIXS line shapes with the exact two- and four-spinon dynamical structure factor of the spin-1/2 Heisenberg chain. In Fig. 3b, we show the fits for three selected momentum transfer values using the exact two- and four-spinon dynamical structure factor,  $S(\mathbf{q}, \omega)$  (where  $\mathbf{q}$  and  $\omega$  stand for the momentum and energy of the created excitation, respectively) in the representation of ref. 22. The obtained exchange coupling,  $J = 249$  meV, is in very good agreement with the value obtained from inelastic neutron scattering data<sup>20</sup>.

Having unambiguously identified the fractionalized spinon excitations in the low-energy sector, we now concentrate on the orbital excitations spectrum in Fig. 4. We find that these are strongly momentum dependent and have a novel, distinct dispersion. This proves that the orbital excitations observed here are of collective nature. The  $xz$  excitation has the largest dispersion, of  $\sim 0.2$  eV, and has a spectrum containing two peculiar components: a lower branch dispersing with periodicity  $\pi$  and, above that, an incoherent spectrum with a double-oval shape. This spectrum is strikingly similar to seemingly unrelated angle-resolved photoemission spectra of 1D copper oxides, which evidence spin-charge separation (see, for example, fig. 3 of ref. 3). This is an indication that the observed orbital dispersion is related to an analogous separation of degrees of freedom.

To test this conjecture, we derived a microscopic model that describes the spin-orbital interactions in  $\text{Sr}_2\text{CuO}_3$  (Methods). The low-energy Kugel-Khomskii Hamiltonian for this was obtained<sup>8,23</sup>

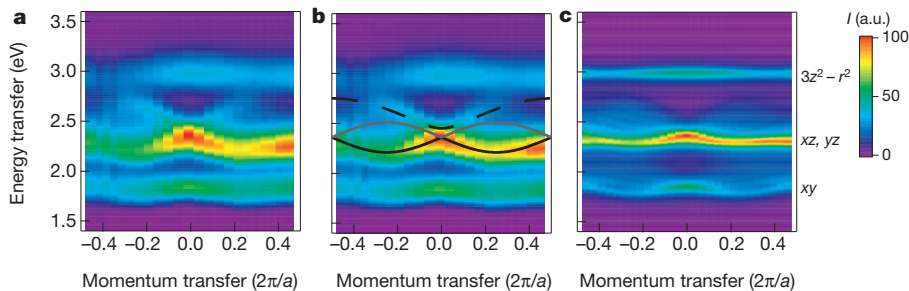
from the charge transfer model of  $\text{Sr}_2\text{CuO}_3$  in ref. 24. The crucial part of the Hamiltonian, responsible for the  $xz$  orbital propagation, is

$$H = -J_O \sum_{j,\sigma} (c_{j\sigma}^\dagger c_{j+1,\sigma} + \text{h.c.}) + J \sum_j \mathbf{S}_j \cdot \mathbf{S}_{j+1} + E_O \sum_j (1 - n_j) \quad (1)$$

where  $J$  and  $J_O$  are respectively the spin and orbital exchange constants, both of which are fixed by the charge transfer model<sup>24</sup>;  $E_O$  is the  $xz$  orbital on-site energy;  $c_{j\sigma}^\dagger$  and  $c_{j\sigma}$  are operators that respectively annihilate and create particles in  $xz$  orbital with spin  $\sigma$ ;  $\mathbf{S}_j$  is the spin of the particle in  $x^2 - y^2$  orbital;  $1 - n_j$  counts the number of particles in  $xz$  orbital; and 'h.c.' denotes Hermitian conjugate. The first term in the Hamiltonian describes the propagation of orbital excitations through the lattice. The second term represents the usual Heisenberg interaction between spins, which vanishes on bonds where an orbital excitation is present (Methods).

From purely a mathematical standpoint, equation (1) is identical to the 1D  $t$ - $J$  model—describing the hopping of holes doped into an antiferromagnetic spin chain in which the spins interact via superexchange  $J$ —with the orbital superexchange,  $J_O$ , taking the place of  $t$ , the hole-hopping amplitude in the  $t$ - $J$  model. This implies that the propagation of a single orbital excitation in the  $J_O$ - $J$  model above is equivalent to the propagation of a single hole in the 1D  $t$ - $J$  Hamiltonian. Because in the 1D  $t$ - $J$  model a hole breaks up into a free spinon and holon, in the  $J_O$ - $J$  model the orbital excitation will separate into a free spinon and orbiton.

Before calculating RIXS spectral functions and quantifying the separation of spin and orbital degrees of freedom in  $\text{Sr}_2\text{CuO}_3$ , it is instructive to consider the overall features of the momentum-dependent RIXS spectra within a slave-boson picture in terms of an orbital-spin separation ansatz. This ansatz is analogous to the mean-field slave-boson picture for charge-spin separation<sup>25,26</sup>, which stipulates that the hole spectral function is given by the convolution of the spectral function of a free holon and a free spinon. In orbital-spin separation ansatz, free orbitons and spinons, created by the operators  $o_i^\dagger$  and  $s_{i\sigma}^\dagger$ , form the particle creation operator  $c_{i\sigma}^\dagger = s_{i\sigma}^\dagger o_i^\dagger$ . On applying this formalism to the Hamiltonian in equation (1), the energies of non-interacting orbitons and spinons with momentum  $k$  become  $\varepsilon_O(k) = E_O - 2J_O \cos(k)$  and  $\varepsilon_S(k) = -J \cos(k)$ , respectively. In this ansatz, the spectral function for the orbital excitation produced in RIXS,  $c_{q\sigma}^\dagger = \sum_k o_{q-k} s_{\sigma k}^\dagger$  (where  $q$ , the  $x$  component of  $\mathbf{q}$ , is here the momentum transfer along the  $\text{CuO}_3$  chain direction), is determined by a convolution of these two dispersions. The resulting excitation continuum is shown in Fig. 4b and has three defining features<sup>25,26</sup>: a prominent lower edge with energy,  $E(q) = E_O - 2J_O |\sin(q)|$ , that is entirely determined by the dispersion of the orbiton; a distinctive upper orbiton edge at  $E(q) = E_O + 2J_O |\sin(q)|$ ; and a broader upper limit of the orbiton-spinon continuum at  $F(q) = E_O + \sqrt{J^2 + 4J_O^2 - 4JJ_O \cos(q)}$ . These dispersion curves are directly compared to our RIXS experiments in



**Figure 4 | Dispersion of orbital excitations: comparison between experiment and *ab initio* calculations.** **a**, **c**, Comparison between the orbital region in Fig. 1c (a, left) and theory, using the  $J_O$ - $J$  model (c, right). **b**, Lower and upper edges of the orbiton dispersion (solid black and grey lines) and the

upper edge of the spinon-orbiton continuum (dashed line), as calculated using the orbital-spin separation ansatz. See main text and Supplementary Information for further details.



Fig. 4b. The measured spectral weight that accumulates at the first two lower-energy features does so with periodicity  $\pi$ , which is a direct consequence of spin–orbital separation.

To quantify the spectral weights within the orbiton–spinon continuum, we have calculated for the  $J_0$ – $J$  Hamiltonian the orbital excitation Green’s function by exactly diagonalizing the Hamiltonian on 28 lattice sites (as in the spin–charge separation studies<sup>3</sup>, finite-size effects are negligible (they are estimated at  $\sim 0.01$  eV)). From this, we calculate the RIXS spectrum following ref. 13, that is, by expressing the RIXS amplitude as a product of the single-ion local RIXS effective operator and the orbital excitation Green’s function (Methods).

We find excellent agreement between theory and experiment (Fig. 4): both the sine-like  $xz$  orbiton dispersion and the  $xz$  spinon–orbiton continuum (the ‘double-oval’ incoherent spectrum), which are the hallmarks of spin–orbital separation, are present in the theory. The calculations also show that, in contrast to the  $xz$  orbital, the  $xy$  orbital has a small dispersion and that the excitations of both the  $3z^2 - r^2$  orbital and the  $yz$  orbital are dispersionless, as is observed experimentally. This is an independent merit of the model because no fitting of dispersions to experimental data is involved.

The large orbiton dispersion observed in this study is the key feature that distinguishes  $\text{Sr}_2\text{CuO}_3$  from other systems with orbital excitations<sup>27–29</sup>, and relies on the 1D character of  $\text{Sr}_2\text{CuO}_3$ . In a system of higher dimensionality, orbitons interact with magnetic excitations, which tend to slow them down and thus reduce their dispersion. In one dimension, orbitons can avoid these renormalization effects by means of spin–orbital separation.

## METHODS SUMMARY

We applied the technique of high-resolution RIXS with the incident photon energy tuned to the  $L_3$  edge ( $2p_{3/2} \rightarrow 3d$  resonance) of the copper ion (around 931 eV) and a total experimental resolution of 140 meV. The experiments were performed at the ADRESS beamline of the Swiss Light Source at the Paul Scherrer Institut<sup>17,18</sup>. The RIXS spectrometer was located at a fixed scattering angle of either  $\Psi = 90^\circ$  or  $130^\circ$  and was collecting signal within the solid angle of  $(19.3 \times 3.3) \text{ mrad}^2$  (horizontal  $\times$  vertical detection direction; for a sketch of the experimental geometry, see Supplementary Fig. 1).

Single-crystal samples of  $\text{Sr}_2\text{CuO}_3$  were grown by the floating-zone method and freshly cleaved before the RIXS experiment. The surface normal to the sample, [010], and the propagation direction of the chains, [100], were oriented parallel to the scattering plane. The sample was cooled with a helium-flow cryostat to 14 K during the measurements. Incident photons were linearly polarized in the scattering plane ( $\pi$ -orientation). The momentum transfer along the chains,  $q$ , was varied by changing the incidence angle in steps of  $5 \pm 1^\circ$ .

Theoretical modelling was done by deriving a Kugel’–Khomskii Hamiltonian from the charge transfer model and then writing it in a representation similar to the one used for the  $t$ – $J$  model. Using this model, we calculated the spectral functions for the case of a single orbital excitation in the antiferromagnetic background by solving the Hamiltonian numerically on 28 sites. Such calculations were repeated separately for the two distinct dispersive orbital excitations, whereas for the non-dispersive excitations we confirmed that the parameters of the model did not allow for a dispersion larger than the experimental resolution. Finally, the RIXS cross-section was calculated by multiplying the calculated spectral functions for the orbital excitations by the local RIXS form factors.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 28 September 2011; accepted 17 February 2012.

Published online 18 April; corrected 2 May 2012 (see full-text HTML version for details).

1. Kugel’, K. I. & Khomskii, D. I. The Jahn–Teller effect and magnetism: transition metal compounds. *Sov. Phys. Usp.* **25**, 231–256 (1982).
2. Giamarchi, T. *Quantum Physics in One Dimension* (Clarendon Press, 2004).
3. Kim, C. *et al.* Observation of spin-charge separation in one-dimensional  $\text{SrCuO}_2$ . *Phys. Rev. Lett.* **77**, 4054–4057 (1996).
4. Fujisawa, H. *et al.* Angle-resolved photoemission study of  $\text{Sr}_2\text{CuO}_3$ . *Phys. Rev. B* **59**, 7358–7361 (1999).
5. Kim, B. J. *et al.* Distinct spinon and holon dispersions in photoemission spectral functions from one-dimensional  $\text{SrCuO}_2$ . *Nature Phys.* **2**, 397–401 (2006).

6. Anderson, P. W. The resonating valence bond state in  $\text{La}_2\text{CuO}_4$  and superconductivity. *Science* **235**, 1196–1198 (1987).
7. Senthil, T. & Fisher, M. P. A. Fractionalization in the cuprates: detecting the topological order. *Phys. Rev. Lett.* **86**, 292–295 (2001).
8. Wohlfeld, K., Daghofer, M., Nishimoto, S., Khaliullin, G. & van den Brink, J. Intrinsic coupling of orbital excitations to spin fluctuations in Mott insulators. *Phys. Rev. Lett.* **107**, 147201 (2011).
9. Ament, L., van Veenendaal, M., Devereaux, T., Hill, J. P. & van den Brink, J. Resonant inelastic X-ray scattering studies of elementary excitations. *Rev. Mod. Phys.* **83**, 705–767 (2011).
10. Moretti Sala, M. *et al.* Energy and symmetry of dd excitations in undoped layered cuprates measured by  $\text{Cu } L_3$  resonant inelastic x-ray scattering. *N. J. Phys.* **13**, 043026 (2011).
11. Ament, L. J. P., Ghiringhelli, G., Moretti Sala, M., Braicovich, L., & van den Brink, J. Theoretical demonstration of how the dispersion of magnetic excitations in cuprate compounds can be determined using resonant inelastic x-ray scattering. *Phys. Rev. Lett.* **103**, 117003 (2009).
12. Luo, J., Trammell, G. T. & Hannon, J. P. Scattering operator for elastic and inelastic resonant x-ray scattering. *Phys. Rev. Lett.* **71**, 287–290 (1993).
13. Haverkort, M. W. Theory of resonant inelastic X-ray scattering by collective magnetic excitations. *Phys. Rev. Lett.* **105**, 167404 (2010).
14. Schlappa, J. *et al.* Collective magnetic excitations in the spin ladder  $\text{Sr}_{14}\text{Cu}_{24}\text{O}_{41}$  Measured using high-resolution resonant inelastic X-Ray scattering. *Phys. Rev. Lett.* **103**, 047401 (2009).
15. Braicovich, L. *et al.* Magnetic excitations and phase separation in the underdoped  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$  superconductor measured by resonant inelastic X-ray scattering. *Phys. Rev. Lett.* **104**, 077002 (2010).
16. Le Tacon, M. *et al.* Intense paramagnon excitations in a large family of high-temperature superconductors. *Nature Phys.* **7**, 725–730 (2011).
17. Strocov, V. N. *et al.* High-resolution soft X-ray beamline ADRESS at the Swiss Light Source for resonant inelastic X-ray scattering and angle-resolved photoelectron spectroscopies. *J. Synchrotron Radiat.* **17**, 631–643 (2010).
18. Ghiringhelli, G. *et al.* SAXES, a high resolution spectrometer for resonant x-ray emission in the 400–1600 eV energy range. *Rev. Sci. Instrum.* **77**, 113108 (2006).
19. Hozoi, L., Siurakshina, L., Fulde, P. & van den Brink, J. Ab initio determination of Cu 3d orbital energies in layered copper oxides. *Sci. Rep.* **1**, 1–4 (2011).
20. Walters, A. C. *et al.* Effect of covalent bonding on magnetism and the missing neutron intensity in copper oxide compounds. *Nature Phys.* **5**, 867–872 (2009).
21. Glawion, S. *et al.* Two-spinon and orbital excitations of the spin-Peierls system  $\text{TiOCl}$ . *Phys. Rev. Lett.* **107**, 107402 (2011).
22. Caux, J.-S. & Hagemans, R. The four-spinon dynamical structure factor of the Heisenberg chain. *J. Stat. Mech.* **2006**, P12013 (2006).
23. Oleś, A. M., Khaliullin, G., Horsch, P. & Feiner, L. F. Fingerprints of spin-orbital physics in cubic Mott insulators: magnetic exchange interactions and optical spectral weights. *Phys. Rev. B* **72**, 214431 (2005).
24. Neudert, R. *et al.* Four-band extended Hubbard Hamiltonian for the one-dimensional cuprate  $\text{Sr}_2\text{CuO}_3$ : distribution of oxygen holes and its relation to strong intersite Coulomb interaction. *Phys. Rev. B* **62**, 10752–10765 (2000).
25. Brunner, M., Assaad, F. F. & Muramatsu, A. Single hole dynamics in the one dimensional t-J model. *Eur. Phys. J. B* **16**, 209–212 (2000).
26. Suzuura, H. & Nagaosa, N. Spin-charge separation in angle-resolved photoemission spectra. *Phys. Rev. B* **56**, 3548–3551 (1997).
27. Macfarlane, R. M. & Allen, J. W. Exciton bands in antiferromagnetic  $\text{Cr}_2\text{O}_3$ . *Phys. Rev. B* **4**, 3054–3067 (1971).
28. Grüninger, M. *et al.* Experimental quest for orbital waves. *Nature* **418**, 39–40 (2002).
29. Ulrich, C. *et al.* Momentum dependence of orbital excitations in Mott-insulating titanates. *Phys. Rev. Lett.* **103**, 107205 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was performed at the ADRESS beamline of the Swiss Light Source using the SAXES instrument jointly built by the Paul Scherrer Institut, Switzerland, and Politecnico di Milano, Italy. We acknowledge support from the Swiss National Science Foundation and its NCCR MaNEP. K.W. acknowledges support from the Alexander von Humboldt foundation and discussions with M. Daghofer and S.-L. Drechsler. J.-S.C. acknowledges support from the Foundation for Fundamental Research on Matter and from the Netherlands Organisation for Scientific Research. S.S. and A.R. acknowledge the support of the European contract NOVMAQ. This research benefited from the RIXS collaboration supported by the Computational Materials Science Network programme of the Division of Materials Science and Engineering, US Department of Energy, grant no. DE-SC0007091.

**Author Contributions** J.S., T.S. and H.M.R. planned the experiment. S.S. and A.R. fabricated the samples. J.S., K.J.Z., V.N.S. and T.S. carried out the experiment. J.S. and M.M. carried out the data analysis. C.M. helped with the data analysis. K.W. and J.v.d.B. developed the theory for the spin–orbital separation with assistance from M.W.H., L.H. and S.N. J.-S.C. provided the theory for the spin excitations. J.S., K.W., K.J.Z., H.M.R., J.v.d.B. and T.S. wrote the paper with contributions from all co-authors. L.P., H.M.R., J.v.d.B. and T.S. supervised the project.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.S. (justine.schlappa@helmholtz-berlin.de) or T.S. (thorsten.schmitt@psi.ch).

## METHODS

**Experiment.** We applied the technique of high-resolution RIXS with the incident photon energy tuned to the  $L_3$  edge ( $2p_{3/2} \rightarrow 3d$  resonance) of the copper ion (around 931 eV) and a total experimental resolution of 140 meV. The experiments were performed at the ADDRESS beamline of the Swiss Light Source at the Paul Scherrer Institut<sup>17,18</sup>. Single-crystal samples of  $\text{Sr}_2\text{CuO}_3$  were grown by the floating-zone method and freshly cleaved before the experiment. The samples were mounted such that the surface normal, [010], and the chain propagation direction, [100], were oriented parallel to the scattering plane (Supplementary Information, section 4, and Supplementary Fig. 1). The polarization vector of the incident light,  $\epsilon_{\text{in}}$ , was parallel to the scattering plane ( $\pi$ -orientation). This yielded the maximum cross-section for the copper  $2p^6 3d^9 \rightarrow 2p^5 3d^{10}$  transition. The RIXS spectrometer was located at a fixed scattering angle of either  $\Psi = 90^\circ$  or  $130^\circ$ , collecting signal within a solid angle of  $(19.3 \times 3.3) \text{ mrad}^2$  (horizontal  $\times$  vertical detection direction). The momentum transfer along the chains,  $q$ , was varied by changing the incidence angle in steps of  $5 \pm 1^\circ$ . This results in the RIXS signal being polarization dependent and the relative intensity of the orbital excitations varying as a function of  $q$ . (For details on the polarization dependence of the orbital excitations, see the discussion of the RIXS cross-section below.) Supplementary Fig. 2a shows the RIXS line spectra, which correspond to the intensity map data in Fig. 1c, and Supplementary Fig. 2b shows the variation in the polarization of the incident photons with  $q$ . We estimate the systematic errors in the momentum transfer to be about  $0.006 \times 2\pi/a$  and  $0.013 \times 2\pi/a$  at the edge and the centre of the Brillouin zone, respectively. The error in determining the energy zero of the energy transfer scale is about 30 meV. The sample was cooled with a helium-flow cryostat to 14 K during the measurements.

**Theory of spin excitations.** Owing to the particular choice of the experimental geometry, the RIXS cross-section for spin excitations (that is, the magnetic part of the spectrum extending up to  $\sim 0.8 \text{ eV}$ ) is directly proportional to the two-point dynamical spin–spin correlation function (local RIXS form factors are independent of transferred momentum here); see refs 9, 11–14. Therefore, the magnetic part of the RIXS spectrum was fitted to the two-point dynamical spin–spin correlation function (see below), convolved with a Gaussian to account for experimental resolution and adding a low-energy Gaussian covering the elastic and phonon peak around 50 meV. This captured most of the measured intensity with nearly constant  $J$ , amplitude and resolution. Fixing these three parameters to their mean values gives reasonable fits for all  $q$ . The time-normalized data sets were corrected for the effective scattering volume,  $V_{\text{eff}}$ :

$$\frac{1}{V_{\text{eff}}} = 1 + \frac{\mu_1(\omega)}{\mu_2(\omega)} \frac{\sin(\theta)}{\sin(\beta)}(q)$$

Here  $(\sin(\theta)/\sin(\beta))(q)$  is a geometric term,  $\theta$  and  $\beta$  being the angles between the sample surface and the direction of photon incidence and, respectively, the direction of photon detection; and  $(\mu_1/\mu_2)(\omega)$  is an outgoing-energy-dependent absorption factor derived from the total fluorescence yield X-ray absorption spectra.

The exact solvability of the Heisenberg spin-1/2 chain is exploited to compute its two-point dynamical spin–spin correlation function. At zero magnetization, this is predominantly carried by two- and four-spinon intermediate states whose exact contributions can be written in the thermodynamic limit in terms of fundamental integrals by using the vertex operator approach<sup>22</sup>. The numerical evaluation of these integrals provides the dynamical structure factor throughout the Brillouin zone, yielding sum-rule saturations of the order of 98%.

**Theory of orbital excitations.** We first calculate the energies of the  $3d$  orbitals in  $\text{Sr}_2\text{CuO}_3$  using the *ab initio* quantum chemistry method (Supplementary Information, section 5). Although this method is not tailored to calculate the dispersion of orbital excitations (for which model calculations are needed) owing to the finite size of the cluster, it identifies from first principles the energies of orbital excitations in  $\text{Sr}_2\text{CuO}_3$  for  $\mathbf{q} = \mathbf{0}$  transferred momentum. These energies have been found to be in close agreement with both optical absorption and RIXS experiments<sup>19</sup>. Owing to the large separation of the  $d$ – $d$  excitations energies (Supplementary Table 1), we can separately model the momentum-dependent RIXS cross-section for each orbital excitation for the case of non-negligible inter-orbital hopping. For this (see the next section), the Kugel–Khomskii Hamiltonian for the case of a single orbital excitation (for example the  $x^2 - y^2 \rightarrow xz$  transition) from a ferro-orbital ground state (that is, one in which all  $x^2 - y^2$  orbitals are occupied) is derived on the basis of the established charge transfer model for  $\text{Sr}_2\text{CuO}_3$ . Similar calculations are performed for the other orbital excitation symmetries (see the next section) and the RIXS cross-section based on the solution to these model calculations is calculated (see the discussion of RIXS cross-section below).

**Model Hamiltonian.** In Mott insulators, the Kugel–Khomskii model is the effective low-energy superexchange model for coupled spin and orbital excitations;

compare with ref. 1. We first consider only an  $xz$  orbital excitation, which may hop between the copper sites in a ‘three-step’ perturbative superexchange process with an energy scale  $\sim t_1 t_2 / 2U$ . First the particle in the excited orbital moves from site  $j$  to the neighbouring site  $j+1$  by hopping with amplitude  $\sim t_2$ ; next an intermediate state, in which two particles are on site  $j+1$  and which has an energetic cost equivalent to the on-site Coulomb repulsion  $\sim U$ , is formed; and finally the particle in the ground-state orbital moves from site  $j+1$  to site  $j$  by hopping with amplitude  $t_1$ . Thus, we obtain

$$H_0 = -J_0 \sum_{j,\sigma} (c_{j\sigma}^\dagger c_{j+1,\sigma} + \text{h.c.}) + E_0 \sum_j (1 - n_j) \quad (2)$$

where  $J_0 = (3R_1 + R_2)2t_1 t_2 / U$  and  $E_0$  is the energy cost of the orbital excitation. In  $J_0$ ,  $R_1 = 1/(1 - 3\eta)$  and  $R_2 = 1/(1 - \eta)$  originate in the multiplet structure of the intermediate states of the superexchange processes and depend on the ratio,  $\eta = J_H/U$ , of the Hund’s exchange to the Coulomb repulsion. Although the above equation is given for the Mott–Hubbard limit of the superexchange model, in our calculation we modified the parameters to account for superexchange processes on oxygen atoms in  $\text{Sr}_2\text{CuO}_3$ . For a standard set of charge transfer parameters<sup>24</sup>, this gives  $J_0 \approx 0.075 \text{ eV}$  without any fitting to the experimental data. The values of  $E_0$  are determined from the RIXS experiment and are within 10% of the on-site orbital energies obtained using *ab initio* quantum chemistry cluster calculations<sup>19</sup> for four  $\text{CuO}_4$  plaquettes in  $\text{Sr}_2\text{CuO}_3$  (Supplementary Information, section 5).

In equation (2),  $c_{j\sigma}$  is a fermion operator, acting in the restricted Hilbert space with no double occupancies, that creates an orbital excitation (a hole in the spin background in the language of the  $t$ – $J$  model) at site  $j$  with spin  $\sigma$ , and  $1 - n_j = 1 - \sum_\sigma c_{j\sigma}^\dagger c_{j\sigma}$  is the number operator that counts the number of orbital excitations in the chain. Note that this ‘fermionization’ of the 1D problem was performed by replacing the standard orbital pseudospin operators<sup>23</sup> with fermions by Jordan–Wigner transformation (see ref. 8 for details).

The second part of the Hamiltonian in equation (1) corresponds to the spin dynamics on the bonds where the orbital excitation is not present:

$$H_S = J \sum_j \mathbf{S}_j \cdot \mathbf{S}_{j+1}$$

Here  $J = 4t_1^2/U$  is now the well-known Heisenberg antiferromagnetic superexchange constant. In sum,  $H_S$  and  $H_0$  constitute the Hamiltonian  $H$  from equation (1).

Excitations from the ground-state  $3d x^2 - y^2$  orbital to either the  $xy$  orbital or the  $xz$  orbital are described by  $H$ , but with different effective parameters because the hopping amplitudes from each of these orbitals to the neighbouring bonding oxygen orbitals and, subsequently, to the neighbouring copper sites differ. In particular, for the  $xy$  orbital the effective hopping amplitude along the chain is  $\sim 25\%$  smaller than that for the  $xz$  orbital owing to the formation of bonding and antibonding states with  $2p$  orbitals on oxygen atoms outside the copper–oxygen chain. For orbital excitations involving the  $yz$  and  $3z^2 - r^2$  orbitals, the dispersion vanishes because the hopping matrix elements to the neighbouring oxygen orbitals along the chain are either much smaller than for the  $xy$  or  $xz$  orbital excitations ( $3z^2 - r^2$  orbital) or are vanishingly small<sup>24</sup> ( $yz$  orbital).

**RIXS cross-section.** Local effective RIXS operators have been derived for the  $\text{Cu}^{2+}$  ( $3d^9$ ) electronic configuration<sup>12,30</sup> (see also Supplementary Information, section 3). However, several simplifications beyond the local effective scattering operators as used in refs 12–13, 30 are justified, because the spin of the particle in the excited orbital is to a good approximation conserved (Hund’s exchange is one order of magnitude smaller than the on-site Coulomb repulsion<sup>23,24</sup> during orbiton propagation). First, the spectrum obtained is independent of the spin of the particle in the excited orbital. Second, the spin and the orbital character of the excited particle are conserved during propagation. The RIXS intensity for an orbital excitation is therefore given by the sum of the intensities for scattering from initial states with spin up or down to final states with the spin either flipped or conserved. This allows us to express the RIXS intensity as the product<sup>11,13</sup> of polarization-dependent intensities and the non-local dynamical structure factor  $O(q, \omega)$ , which is defined as the spectral function of the single orbital excitation created using operator  $c_{j\sigma}$  and propagating with Hamiltonian  $H$  (ref. 8). This structure factor is either equal to one, for ‘non-dispersive’ orbital excitations ( $3z^2 - r^2$  and  $yz$  orbitals; see previous section), or is calculated numerically by exactly solving the Hamiltonian  $H$  on a finite chain of 28 sites. The latter is done separately for  $xz$  and  $xy$  orbital excitations. We note that such a separation is possible because orbiton–orbiton interactions vanish for one-orbiton excitations in the chain. The results of these calculations are presented in Fig. 4c.

30. van Veenendaal, M. Polarization dependence of  $L$ - and  $M$ -edge resonant inelastic X-ray scattering in transition-metal compound. *Phys. Rev. Lett.* **96**, 117404 (2006).

# Liquid-crystal-mediated self-assembly at nanodroplet interfaces

J. A. Moreno-Razo<sup>1</sup>, E. J. Sambriski<sup>2</sup>, N. L. Abbott<sup>3</sup>, J. P. Hernández-Ortiz<sup>4</sup> & J. J. de Pablo<sup>5</sup>

**Technological applications of liquid crystals have generally relied on control of molecular orientation at a surface or an interface<sup>1,2</sup>. Such control has been achieved through topography, chemistry and the adsorption of monolayers or surfactants<sup>2,3</sup>. The role of the substrate or interface has been to impart order over visible length scales and to confine the liquid crystal in a device. Here, we report results from a computational study of a liquid-crystal-based system in which the opposite is true: the liquid crystal is used to impart order on the interfacial arrangement of a surfactant. Recent experiments on macroscopic interfaces have hinted that an interfacial coupling between bulk liquid crystal and surfactant can lead to a two-dimensional phase separation of the surfactant at the interface<sup>4</sup>, but have not had the resolution to measure the structure of the resulting phases. To enhance that coupling, we consider the limit of nanodroplets, the interfaces of which are decorated with surfactant molecules that promote local perpendicular orientation of mesogens within the droplet. In the absence of surfactant, mesogens at the interface are all parallel to that interface. As the droplet is cooled, the mesogens undergo a transition from a disordered (isotropic) to an ordered (nematic or smectic) liquid-crystal phase. As this happens, mesogens within the droplet cause a transition of the surfactant at the interface, which forms new ordered nanophases with morphologies dependent on surfactant concentration. Such nanophases are reminiscent of those encountered in block copolymers<sup>5</sup>, and include circular, striped and worm-like patterns.**

Liquid crystals (LCs) exhibit some features of crystalline solids, such as long-range orientational order, and some features of liquids, such as the ability to flow<sup>6,7</sup>. Their mesomorphic behaviour arises from anisotropic molecular shapes and interactions that produce characteristic, non-isotropic spatial arrangements. A variety of LC phases can be obtained by changing density or temperature, by adding co-solvents, or by mixing different molecules<sup>6</sup>. Here we show that, when confined to small droplets, liquid crystals can induce the formation of ordered domains of surfactants localized at the droplet interfaces. This is in contrast to past research, which has largely elucidated the effects that substrates or interfaces have on the long-range orientation of liquid crystals. The realization that LCs can direct the formation of ordered nanostructured morphologies at fluid interfaces enables new hierarchical assembly strategies.

LCs are remarkably sensitive to interfacial chemistry and topography. Distinct interfacial events that occur on the scale of nanometres can trigger a response of the LC that is amplified over macroscopic length scales, giving rise to optical textures that can be examined directly by optical microscopy<sup>3,8–10</sup>. The converse phenomenon, that of LC order influencing interfacial structure, has received little attention. Recent reports, however, suggest that surfactants at an LC–water interface are not distributed homogeneously, but can form macroscopic domains that are visible to the naked eye<sup>4</sup>. In related work, researchers have also

examined the two-dimensional phases of fatty-acid monolayers at the interface of a liquid crystal with water<sup>11</sup>. The question that thus arises is whether LC order can be used to drive the organization and even possibly induce the self-assembly of surface-active molecules at an interface.

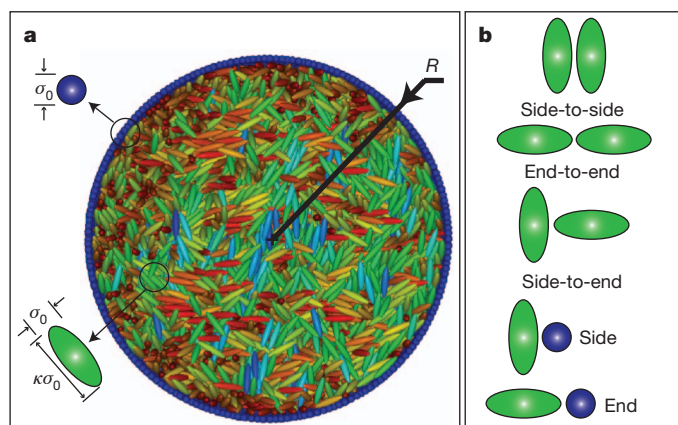
Nematic LCs exhibit mechanical constants that penalize splay, bending and twist deformations. As the dimensions of an LC sample become smaller and reach length scales commensurate with characteristic mechanical correlation lengths, it becomes increasingly difficult to deform the material. Therefore, to maximally enhance the effect of LC order on an interface, we focus on LC nanodroplets, which can be prepared experimentally by simple emulsification<sup>10</sup>, or by using more elaborate techniques involving microfluidic methods or polyelectrolyte multilayers<sup>12–14</sup>. From a technological point of view, interest in water-in-LC and LC-in-water emulsions has been fuelled by recent demonstrations of transparent nematic phases<sup>15</sup>, photonic switches<sup>16</sup>, microresonators<sup>17,18</sup>, and bacterial biosensors<sup>19</sup>.

Here we use a minimalist model to identify key characteristic features and map out system behaviour. The model includes the main attributes of experimentally realizable systems, but omits complex interactions that might preclude an unambiguous interpretation of its predictions. A schematic representation of the system is provided in Fig. 1 (see also Supplementary Fig. 1). Mesogens are represented as ellipsoids confined to a spherical droplet. The surfactant and solvent molecules at the interface of the droplet are represented by spherical sites all of the same size. The interactions between solvent and surfactant molecules are identical and are limited to excluded volume effects. Thus, any morphological transitions of interfacial solvent and surfactant molecules at the interface can arise only from an indirect interaction mediated by the LC in the droplet. Surfactant molecules promote a perpendicular orientation of the mesogens, whereas water molecules favour a parallel orientation. The concentration (number fraction) of surfactant at the interface is given by  $x_{\text{surf}}$ . Monte Carlo simulations in the microcanonical ensemble are used to determine the spatial order of the LC and surfactant as a function of  $x_{\text{surf}}$ . For convenience, all quantities are reported in reduced units, with respect to a molecular length scale  $\sigma_0$  (the width of the LC ellipsoids, of the order of 1 nm) and a molecular energy scale  $\varepsilon_0$  (of the order of  $\varepsilon_0/k_B = 4,302$  K) such that  $\rho^* = \rho\sigma_0^3$ ,  $U^* = U/\varepsilon_0$ , and  $T^* = k_B T/\varepsilon_0$ , with  $k_B$  being the Boltzmann constant). Additional details pertaining to the model and the simulations are provided in the Supplementary Information. Parameters for the mesogen–mesogen interaction potential are taken from those previously determined for *p*-terphenyl<sup>20</sup>. The droplets considered here have a diameter of  $50\sigma_0$ , which corresponds to approximately 50 nm. A reduced temperature is defined by  $T^* = k_B T/\varepsilon_0$ . In the bulk, the mesogens considered here undergo an isotropic–nematic transition at  $T^* = 1.4$ .

Supplementary Fig. 2 shows the phase diagram for droplets having a homogeneous (pure) surface as a function of anchoring strength. The

<sup>1</sup>Departamento de Física, Universidad Autónoma Metropolitana-Iztapalapa, Apartado Postal 55-534, México 09340, Distrito Federal, México. <sup>2</sup>Department of Chemistry and Biochemistry Delaware Valley College, Doylestown, Pennsylvania 18901, USA. <sup>3</sup>Department of Chemical and Biological Engineering University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. <sup>4</sup>Departamento de Materiales, Universidad Nacional de Colombia, Sede Medellín, Bloque M3-050, Medellín, Colombia. <sup>5</sup>Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.



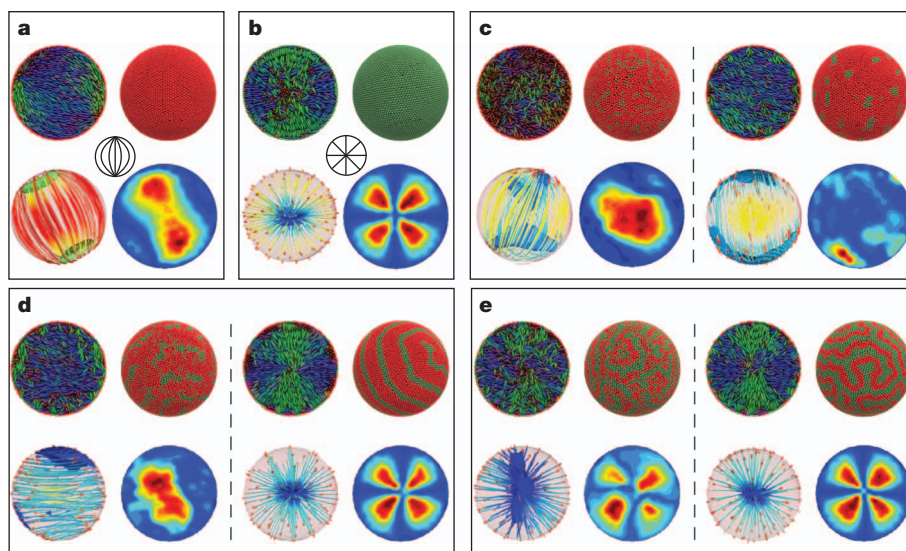


**Figure 1 | Model of the nematic LC droplet.** **a**, Schematic cross-section displaying various interactions arising in an LC droplet of radius  $R = 25\sigma_0$ . Solvent particles (blue spheres) constitute the droplet boundary (consisting of  $n_s = 10,832$  spheres). Mesogens (ovals, coloured according to their respective orientation) are enclosed within the boundary (totalling  $n_m = 13,500$  mesogens). The number density of mesogens is  $\rho^* = n_m / [(4/3)\pi(R/\sigma_0)^3] = 0.206$ . Interactions between boundary particles are governed by  $U_{LJ}$ . Two types of mesogen–boundary interactions are considered: a mesogen–solvent interaction promotes a parallel orientation at the boundary, while a mesogen–surfactant interaction promotes a radial orientation at the boundary. These interactions are governed by  $U_{RS}$ . Mesogen–mesogen interactions are controlled through  $U_{GB}$ . The upper inset shows the spatial dimensions of solvent–surfactant particles (blue), with diameter  $\sigma_0$ . The lower inset shows the spatial dimensions of a prolate mesogen (green oval), with a minor axis of length  $\sigma_0$  and a major axis of length  $\kappa\sigma_0$  (with  $\kappa = 4.4$ ). **b**, Conventions to denote limiting cases of molecular orientations in mesogen–mesogen and mesogen–boundary interactions.

droplet morphologies are generally consistent with those observed in experiments and continuum descriptions of large drops. For the study of surfactant-laden droplets, we focus on the regime corresponding to intermediate anchoring between the mesogens and the surfactant and water particles at the interface. Results from simulations over a wide range of conditions are shown in Figs 2 and 3. The limits  $x_{\text{surf}} = (0.00,$

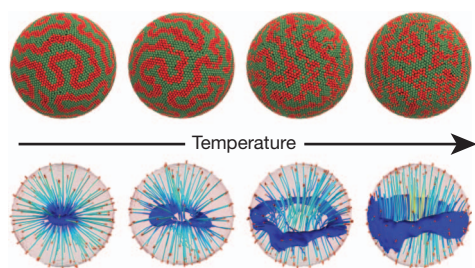
1.00) are shown in Fig. 2a and b, respectively, along with line drawings of the corresponding defects observed at these limits. At high temperatures ( $T^* > 13$ ), the LC structure in the droplet is isotropic. At intermediate temperatures ( $2 < T^* < 13$ ), the mesogens adopt a nematic LC phase. The defect when  $x_{\text{surf}} = 0.00$  is bipolar and the defect when  $x_{\text{surf}} = 1.00$  is radial. Results for intermediate values of  $x_{\text{surf}}$  are shown in Fig. 2c, d and e. When mesogens in the droplet are disordered, the surfactant at the interface is distributed homogeneously throughout the interface, regardless of composition. As the temperature is lowered, mesogens within the droplet undergo a phase transition and enter a nematic phase. The isotropic–nematic transition is accompanied by a morphological transformation of the surfactant at the interface, which in turn adopts highly ordered arrangements. Depending on the concentration of surfactant, the interface exhibits circular domains, stripes or more intricate patterns, such as those observed for  $x_{\text{surf}} = 0.50$  (Fig. 2e).

In the molecular model considered here, the organization of the surfactant is driven by the orientation and packing of the LC in the immediate vicinity of the interface, where it is energetically favourable for the mesogens to form ordered domains of homeotropic or parallel orientation, depending on whether the nearby concentration of surfactant is high or low, respectively. The order at the interface must then be balanced with that in the bulk of the LC in the droplet, giving rise to the formation of defects such as those shown in Supplementary Fig. 4. To determine the origin of those patterns, we resort to a continuum representation of the LC in terms of short-range, elastic and surface contributions to the free energy. The details of the model are provided in the Supplementary Information. The results of our continuum analysis, summarized in Supplementary Figs 5 and 6, reveal that as the droplet size becomes small, a characteristic length scale arises as a result of the balance between the elastic free energy, which increases as the size of the domains and the underlying defects increases, and the surface free energy, which decreases as the size of the domains increases (see Supplementary Figs 6 and 7). As the domains become small, they reach a length scale below which the LC can no longer accommodate the sharp deformations required to satisfy surface anchoring, thereby leading to sharp increases of surface free energy. Short-range contributions to the free energy are independent of



**Figure 2 | Representative configurations of nanodroplets.** Results are shown for  $x_{\text{surf}} = 0.0$  (no surfactant) and  $x_{\text{surf}} = 1.0$  (complete surfactant coverage) at a temperature  $T^* = 2$  (**a** and **b**). Line drawings in the centre of each panel depict the bipolar (in **a**) and radial (in **b**) defects in the LC droplets. Each panel shows four views: upper left, cross-section showing mesogens; upper right, distribution of water (red) or surfactant (green) molecules at the interface; lower left, contour lines of the director; and lower right, defect core and

polarized radiation transmission. The mesogens in the droplets are assigned different colours according to their specific orientation. Intermediate concentrations of surfactant on the LC droplet surface are shown in **c–e**, using the same representations as described for **a** and **b**. **c**,  $x_{\text{surf}} = 0.10$ ,  $T^* = 13$  (left, isotropic phase) and  $T^* = 8$  (right, nematic phase). **d**,  $x_{\text{surf}} = 0.25$ ,  $T^* = 11$  (left, nematic phase) and  $T^* = 2$  (right, smectic phase). **e**,  $x_{\text{surf}} = 0.50$ ,  $T^* = 11$  (left, nematic phase) and  $T^* = 2$  (right, smectic phase).



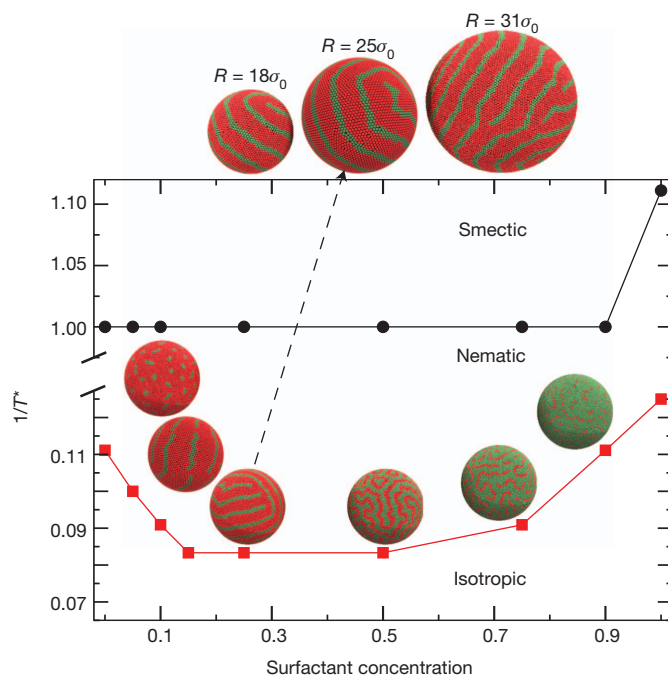
**Figure 3 | Director contour lines and defect core as a function of temperature, corresponding to a surfactant concentration of  $x_{\text{surf}} = 0.5$ .** The four temperatures shown here are  $T^* = 2, 5, 10$  and  $11$ , respectively. At  $T^* = 2$ , a smectic phase is observed. All other temperatures correspond to a nematic phase. As the temperature is lowered, the defect becomes increasingly localized at the centre of the droplet and squeezes the director until the overall configuration of the LC is radial. Ordering of surface particles becomes increasingly pronounced as the temperature is reduced and the LC becomes more ordered.

domain size; the sum of elastic and surface free energy therefore exhibits a distinct minimum, which becomes more pronounced as the droplets become smaller. For strong anchoring and large droplets, on the order of  $R = 800$  nm, the minimum disappears.

The structure of the LC within surfactant-laden nanodroplets is different from that expected for bare droplets (in which case, a bipolar structure is observed, as in Fig. 2a). The phase diagram of bare droplets is provided in Supplementary Fig. 2 as a function of anchoring strength. Of the many new structures that arise for inhomogeneous surfaces, a concrete example is shown in Fig. 3 for  $x_{\text{surf}} = 0.50$  (refer to Fig. 2e). In that case, just below the isotropic–nematic transition temperature, the formation of an equatorial line defect appears as a broad band that spans the entire circumference of the droplet. As the temperature is lowered, the defect becomes smaller and ‘squeezes’ the director profile in the droplet. This squeezing process becomes increasingly pronounced as the temperature is lowered, until the defect becomes a small ring that is highly localized at the centre of the droplet. In that limit, the structure of the droplet is clearly radial, and at sufficiently low temperatures, the emergence of a smectic phase is observed.

The surfactant patterns seen in Fig. 2 bear resemblance to the morphologies encountered in diblock copolymers<sup>5</sup>. As shown in Fig. 4, a phase diagram can be constructed so as to provide, at constant droplet size, a taxonomy of surfactant morphologies as a function of concentration. At low concentrations, a circular phase is observed, consisting of regular surfactant domains spread uniformly throughout the surface of the droplet. At higher concentrations, the surfactant forms ordered stripes spaced regularly throughout the interface. Such stripes are reminiscent of the lamellar phase of block copolymers. At even higher concentrations, the surfactant adopts a bicontinuous morphology, in which a single stripe of surfactant extends over the entire surface of the droplet. As the concentration of surfactant is raised further, a worm-like pattern emerges, where the surfactant now surrounds elongated pools of solvent molecules. Finally, at high concentrations, the system adopts an inverted circular phase morphology, in which small pools of solvent are dispersed regularly throughout the interface. The phase diagram presented in Fig. 4 corresponds to droplets of radius  $25\sigma$ , but simulations of smaller ( $R = 18\sigma$ ) and larger ( $R = 31\sigma$ ) systems show similar water-rich and surfactant-rich domains. Figure 4 includes representative lamellar configurations of droplets of varying size for a surfactant concentration of 25%.

Results from Figs 2 and 4 clearly show that a three-dimensional ordering transition of the LC drives a two-dimensional ordering transition of the surfactant at the interface. In the absence of LC, the interfacial surfactant and solvent molecules in our model cannot undergo a phase separation. The morphologies observed in our particle-based simulations were corroborated by results from a well established continuum theory, and suggest that the phenomena put forth here may be



**Figure 4 | Phase diagram showing the ordered patterns adopted by the surfactant at the droplet interface, as a function of temperature and composition.** The formation of circles, stripes, and worm-like conformations are visible at the interface. At low temperatures, LC ordering is smectic. At intermediate temperatures, LC ordering is nematic, while at high temperatures, LC ordering is isotropic. For small droplets, the patterns adopted by the surfactant are relatively insensitive to droplet size, as suggested by the images above the diagram, which correspond to droplet diameters of  $18\sigma_0$ ,  $25\sigma_0$  and  $31\sigma_0$ , respectively.

general. We are not aware of a precedent for such transitions in the experimental literature, which has been largely concerned with droplets in the micrometre range. For such droplets, experimental evidence suggests that surfactants are distributed inhomogeneously at the interface, but past measurements have not had the resolution to discern between distinct morphologies at the interface. For droplets containing simple isotropic liquids, a surfactant at the interface will exhibit a liquid, liquid-expanded, or a gaseous phase, but none of these phases exhibit the degree of organization seen in Figs 2 and 4. The balance between elastic and surface forces, and the high energy premium required to deform the LC in the near vicinity of the interface, are the effects that bring about these new interfacial phases and their characteristic length scale. Collectively, these features seem to be unique to LC materials.

As shown in the Supplementary Information, other types of confinement (for example, cylindrical or planar) also give rise to surfactant domain formation (see Supplementary Fig. 3), but the resulting morphologies do not exhibit the order that is apparent in spherical droplets (and shown in Figs 2 to 4). Also, our calculations were restricted to the range from intermediate to strong anchoring conditions; for very weak anchoring these effects eventually disappear, because the surface energy becomes much weaker than the elastic energy. Surface active molecules or nanoparticles and LCs can be tuned to offer a wide array of anchoring conditions, mechanical constants, and droplet sizes that could, in turn, give rise to even more complex patterns and behaviours than those reported here. Recent advances in nanoparticle synthesis, characterization, and simulation have revealed that surfactants on solid nanoparticles can form ordered structures, and that such nanoparticles can subsequently be assembled into intriguing ordered arrays<sup>21–23</sup>.

The results presented here indicate that surfactants can also form ordered phases on liquid interfaces, provided the underlying liquid is anisotropic. Surfactant molecules in the ordered morphologies



anticipated in this work could in principle be functionalized, thereby providing the basis for hierarchical assembly of more complex superstructures<sup>23</sup>. Our predictions offer a motivation to pursue the synthesis and characterization of LC nanodroplets with carefully controlled interfacial characteristics that might offer new routes for hierarchical assembly of structured materials.

## METHODS SUMMARY

The coarse-grained LC molecules considered in this work are confined to a spherical cavity, whose boundary (interface) consists of spherical particles of radius  $\sigma_0$  and aspect ratio  $\kappa$ , constrained to move on the surface of a sphere of radius  $R$  (ref. 24). The system is shown schematically in Fig. 1. Spherical particles on the droplet surface interact through a Lennard–Jones potential energy function<sup>25,26</sup>. LC molecules interact through a Gay–Berne potential energy function<sup>27</sup>. Preparatory molecular dynamics simulations were performed in the canonical ensemble (at constant number of molecules  $N$ , constant volume  $V$  and constant temperature  $T$ ), followed by extensive equilibration runs in the microcanonical ensemble (constant  $N$ ,  $V$  and constant energy  $E$ ). For most of the simulations the radius of the droplet was held at  $R = 25\sigma_0$ . Additional simulations with smaller ( $R = 18\sigma_0$ ) and larger droplets ( $R = 31\sigma_0$ ) were also carried out to examine finite-size effects. The number of Lennard–Jones particles over the spherical surface was  $N_{\text{LJ}} = 10,832$ , while the number of mesogens was  $N_{\text{GB}} = 13,500$ . The density inside the droplet was  $\rho_{\text{GB}}^* (= \rho\sigma_0^3) = 0.206$ . For aromatic compounds such as *p*-terphenyl such a density corresponds to a value of  $0.70 \text{ g cm}^{-3}$  (ref. 20). All simulations were started at high temperatures  $T^* (= k_{\text{B}}T/\epsilon_0) = 16.0$  in the isotropic phase (Fig. 1 shows a representative configuration). Low-temperature states were attained through gradual cooling. Once a desired temperature was attained, simulations were performed in the *NVE* ensemble. The translational and rotational equations of motion were integrated using the velocity–Verlet algorithm<sup>28</sup>, with a time step  $\delta t^* \left( = \delta t (\epsilon_0/m\sigma_0^2)^{1/2} \right) = 0.001$  (corresponding to about 0.1 ns). Typical runs consisted of at least two million time steps for equilibration, followed by a production period of at least the same duration (equivalent to about 4  $\mu\text{s}$ ). In addition to traditional molecular-dynamics moves, our simulations were supplemented by Monte Carlo moves that consisted of trial exchanges of the identity of surface particles, subject to Metropolis acceptance criteria. Such moves are crucial for equilibration of the surface morphologies presented in this work.

Received 10 October 2011; accepted 19 March 2012.

- Collins, P. J. & Hird, M. *Introduction to Liquid Crystals Chemistry and Physics* (Taylor and Francis, 2004).
- Jerome, B. Surface effects and anchoring in liquid crystals. *Rep. Prog. Phys.* **54**, 391–451 (1991).
- Shah, R. & Abbott, N. L. Principles for measurement of chemical exposure based on recognition-driven anchoring transitions in liquid crystals. *Science* **293**, 1296–1299 (2001).
- Gupta, J., Meli, M., Teren, S. & Abbott, N. L. Elastic energy-driven phase separation of phospholipid monolayers at the nematic liquid-crystal-aqueous interface. *Phys. Rev. Lett.* **100**, 048301 (2008).
- Bates, F. & Fredrickson, G. Block copolymers—designer soft materials. *Phys. Today* **52**, 32–38 (1999).
- de Gennes, P. G. & Prost, J. *The Physics of Liquid Crystals* (Oxford University Press, 1995).
- Barón, M. Definitions of basic terms relating to low-molar-mass and polymer liquid crystals. *Pure Appl. Chem.* **73**, 845–895 (2001).
- Price, A. & Schwartz, D. DNA hybridization-induced reorientation of liquid crystal anchoring at the nematic liquid crystal/aqueous interface. *J. Am. Chem. Soc.* **130**, 8188–8194 (2008).
- Park, J. & Abbott, N. L. Ordering transitions in thermotropic liquid crystals induced by the interfacial assembly and enzymatic processing of oligopeptide amphiphiles. *Adv. Mater.* **20**, 1185–1190 (2008).
- Sivakumar, S., Wark, K., Gupta, J., Abbott, N. L. & Caruso, F. Liquid crystal emulsions as the basis of biological sensors for the optical detection of bacteria and viruses. *Adv. Funct. Mater.* **19**, 2260–2265 (2009).
- Price, A. & Schwartz, D. Fatty-acid monolayers at the nematic/water interface: phases and liquid-crystal alignment. *J. Phys. Chem. B* **111**, 1007–1015 (2007).
- Fernández-Nieves, A. *et al.* Novel defect structures in nematic liquid crystal shells. *Phys. Rev. Lett.* **99**, 157801 (2007).
- Fernández-Nieves, A., Link, D., Marquez, M. & Weitz, D. Topological changes in bipolar nematic droplets under flow. *Phys. Rev. Lett.* **98**, 087801 (2007).
- Gupta, J., Sivakumar, S., Caruso, F. & Abbott, N. Size-dependent ordering of liquid crystals observed in polymeric capsules with micrometer and smaller diameter. *Angew. Chem.* **48**, 1652–1655 (2009).
- Yamamoto, J. & Tanaka, H. Transparent nematic phase in a liquid-crystal-based emulsion. *Nature* **409**, 321–325 (2001).
- Fernández-Nieves, A., Link, D. R. & Weitz, D. A. Polarization dependent Bragg diffraction and electro-optic switching of three-dimensional assemblies of nematic liquid crystal droplets. *Appl. Phys. Lett.* **88**, 121911 (2006).
- Humar, M., Ravník, M., Pajk, S. & Musevic, I. Electrically tunable liquid crystal optical microresonators. *Nature Photon.* **3**, 595–600 (2009).
- Yokohama, H. Tunable whispers. *Nature Photon.* **3**, 560–561 (2009).
- Lin, I. *et al.* Endotoxin-induced structural transformations in liquid crystalline droplets. *Science* **332**, 1297–1300 (2011).
- Luckhurst, G. R. & Simmonds, P. S. J. Computer simulation studies of anisotropic systems. XII. Parameterization of the Gay–Berne potential for model mesogens. *Mol. Phys.* **80**, 233–252 (1993).
- Jackson, A., Myerson, J. & Stellacci, F. Spontaneous assembly of subnanometre-ordered domains in the ligand shell of monolayer-protected nanoparticles. *Nature Mater.* **3**, 330–336 (2004).
- Singh, C. *et al.* Entropy-mediated patterning of surfactant-coated nanoparticles and surfaces. *Phys. Rev. Lett.* **99**, 226106 (2007).
- Kotov, N. & Stellacci, F. Frontiers in nanoparticle research: toward greater complexity of structure and function of nanomaterials. *Adv. Mater.* **20**, 4221–4222 (2008).
- Juffer, A. H. & Berendsen, H. J. C. Dynamical surface boundary conditions: a simple boundary model for molecular dynamics simulations. *Mol. Phys.* **79**, 623–644 (1993).
- Frenkel, D. & Smith, B. *Understanding Molecular Simulations—From Algorithms to Applications* (Academic Press, 1996).
- Allen, M. & Tildesley, D. *Computer Simulation of Liquids* (Oxford Science Publications, 1987).
- Gay, J. G. & Berne, B. J. Modification of the overlap potential to mimic a linear site-site potentials. *J. Chem. Phys.* **74**, 3316–3319 (1981).
- Ilnytskyi, M. & Wilson, M. R. A domain decomposition molecular dynamics program for the simulation of flexible molecules of spherically-symmetrical and nonspherical sites. II. Extension to NVT and NPT ensembles. *Comput. Phys. Commun.* **148**, 43–58 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The continuum analysis of liquid crystal interfaces and the development of the corresponding theory were supported by the Department of Energy, Basic Energy Sciences, Biomaterials Program (DE-SC0004025). The calculations of surfactant organization at nanodroplet interfaces were supported by the National Science Foundation (DMR-1121288). The calculations reported here were performed on computational facilities supported by the National Science Foundation (DMR-1121288).

**Author Contributions** J.A.M.-R. performed the molecular dynamics and Monte Carlo simulations presented in this work. E.J.S. performed critical analysis of droplet structures. N.L.A. was involved in study design. J.P.H.-O. performed the continuum calculations presented in this work. J.J.d.P. designed the study, analysed data, and wrote the paper. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.J.d.P. ([depablo@engr.wisc.edu](mailto:depablo@engr.wisc.edu)).



# A perovskitic lower mantle inferred from high-pressure, high-temperature sound velocity data

Motohiko Murakami<sup>1</sup>, Yasuo Ohishi<sup>2</sup>, Naohisa Hirao<sup>2</sup> & Kei Hirose<sup>3,4</sup>

**The determination of the chemical composition of Earth's lower mantle is a long-standing challenge in earth science. Accurate knowledge of sound velocities in the lower-mantle minerals under relevant high-pressure, high-temperature conditions is essential in constraining the mineralogy and chemical composition using seismological observations<sup>1</sup>, but previous acoustic measurements were limited to a range of low pressures and temperatures. Here we determine the shear-wave velocities for silicate perovskite and ferropericlase under the pressure and temperature conditions of the deep lower mantle using Brillouin scattering spectroscopy<sup>2</sup>. The mineralogical model that provides the best fit to a global seismic velocity profile<sup>1</sup> indicates that perovskite constitutes more than 93 per cent by volume of the lower mantle, which is a much higher proportion than that predicted by the conventional peridotitic mantle model. It suggests that the lower mantle is enriched in silicon relative to the upper mantle, which is consistent with the chondritic Earth model. Such chemical stratification implies layered-mantle convection with limited mass transport between the upper and the lower mantle.**

It is widely accepted, on the basis of petrological evidence, that at least Earth's uppermost mantle, perhaps down to the top of the mantle transition zone, has a peridotitic (pyrolitic) bulk composition<sup>3,4</sup>. However, a variety of chemical compositions, ranging from peridotitic to chondritic, have been proposed for the lower mantle. The correct composition has long remained a matter of debate owing to a lack of conclusive arguments. Because the isochemical phase transitions in  $(\text{Mg,Fe})_2\text{SiO}_4$  well explain the global existence of the seismic discontinuities at depths of  $\sim 410$  and  $\sim 660$  km (ref. 5), the lower mantle has been conventionally been considered to be peridotitic in composition, with a Mg/Si ratio of  $\sim 1.3$ . The same ratio is found in the upper mantle, which is dominated by  $(\text{Mg,Fe})_2\text{SiO}_4$  olivine. However, this Mg/Si ratio is significantly higher than that of chondritic meteorites (Mg/Si  $\approx 1.0$ ), which are usually assumed to be similar in composition to the materials from which Earth formed. The apparent depletion of silicon in the mantle (the 'missing silicon' problem) has provoked much debate that the shortfall is balanced by the presence of silicon in the core<sup>6</sup> or relative silicon enrichment in the lower mantle. Such chemical stratification in the mantle should have occurred during the solidification of the massive magma ocean at the very beginning of Earth's history<sup>7</sup>. Subsequent solid-state convection tends to homogenize the mantle, but the primordial chemical stratification may still be preserved today. This issue is strongly related to the sort of convection that took place in the mantle throughout Earth's history. It has also been proposed that Earth formed preferentially from meteorites with higher Mg/Si ratios owing to the possible radial chemical zonation in the Solar System<sup>8</sup>.

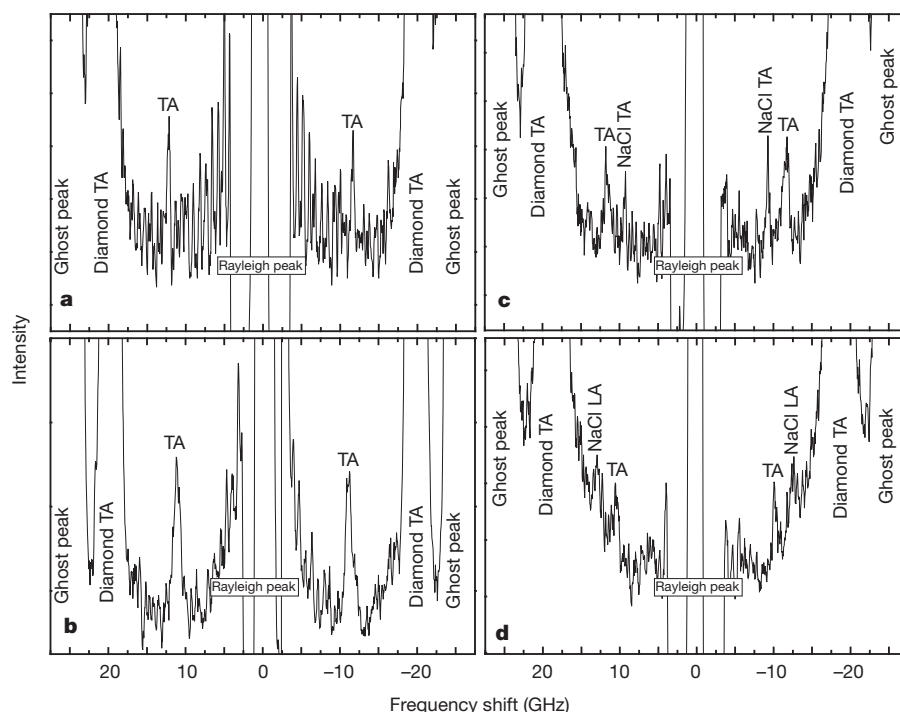
The mineral assemblage of the lower mantle has been examined using density measurements under high-pressure ( $P$ ), high-temperature ( $T$ ) conditions<sup>9,10</sup>. However, recent computer simulations demonstrated that such experimentally derived density and bulk modulus do not

place unique constraints on mantle composition, owing to their intrinsic uncertainties, whereas data on shear velocity ( $V_s$ ) strongly constrain the lower-mantle models<sup>11</sup>. It is thus crucial to obtain the reliable  $V_s$  data on the major lower-mantle constituents, namely silicate perovskite (pv) and  $(\text{Mg,Fe})\text{O}$  ferropericlase (fp), under relevant high-pressure, high-temperature conditions. Recent progress in Brillouin scattering spectroscopy optimized for extreme high-pressure conditions has enabled us to measure  $V_s$  for  $\text{MgSiO}_3$  pv<sup>12</sup>, post-pv<sup>13</sup>, and  $\text{MgO}$  periclase<sup>14</sup> at pressures up to 172 GPa. Although those sound velocity data do provide valuable information on the lower-mantle mineralogy, the effects of chemical impurity and high temperature remain unsettled. Although aluminium impurity was reported to have a strong effect on thermoelastic properties of pv<sup>15</sup>, the sound velocity of aluminium-bearing  $\text{MgSiO}_3$  pv has been measured only below 45 GPa (refs 16, 17). The crossover from high-spin iron to low-spin iron in fp<sup>18</sup> may have an anomalous effect on elasticity at pressures found in the mid-lower mantle<sup>19</sup>, but available sound velocity data for fp, especially for its low-spin state, are still limited. More importantly, sound velocity measurements for pv and fp have never been performed under simultaneous high-pressure and high-temperature conditions corresponding to those in the lower mantle. In this work, we determined  $V_s$  for aluminous silicate pv (Al pv) and fp at pressures up to 124 GPa and at  $T = 300$  K. The high-pressure, high-temperature measurements were also conducted on  $\text{MgSiO}_3$  pv and  $\text{MgO}$  at  $T = 2,700$  K and at pressures up to 91 GPa by using a new Brillouin scattering system<sup>2</sup>.

We obtained very sharp Brillouin peaks from the transverse acoustic modes of each phase over the entire  $P$ - $T$  range we explored (Fig. 1 and Supplementary Tables 1–4). No significant peak broadening was observed with increasing pressure. The results demonstrate slightly lower values of  $V_s$  for  $\text{MgSiO}_3$  (+4 wt%  $\text{Al}_2\text{O}_3$ ) pv than for pure  $\text{MgSiO}_3$  pv<sup>12</sup> at equivalent pressures (Fig. 2a). An anomalous velocity change was observed for  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  fp at around 40–60 GPa (Fig. 2b), which is attributed to the iron spin crossover. A flattening of the velocity up to 50 GPa and then its steep increase to 60 GPa are in excellent agreement with recent Brillouin scattering measurements<sup>20</sup>. However, we did not find evidence for elastic softening at the spin crossover as previously reported on the basis of impulsive stimulated scattering spectroscopy<sup>19</sup>. At pressures greater than 60 GPa,  $V_s$  increased slowly with pressure relative to its behaviour for pure  $\text{MgO}$ .

These  $P$ - $V_s$  profiles of Al pv and fp were fitted to the third-order Eulerian finite-strain equation<sup>21</sup> to obtain the adiabatic shear modulus ( $G_0$ ) and its pressure derivative ( $G_0' = dG/dP$ ). For this regression, the zero-pressure volume and the isothermal bulk modulus and its pressure derivative were adopted from recent experiments<sup>9,17</sup>. We obtained best-fit values of  $G_0 = 166(1)$  GPa and  $G_0' = 1.57(5)$  for Al pv. This  $G_0$  value is consistent with the previous result of 165(2) GPa determined under ambient conditions<sup>16</sup>. The  $G_0'$  value is consistent with the value of 1.56(4) obtained for pure  $\text{MgSiO}_3$  pv<sup>12</sup>, indicating the minimal effect of aluminium on  $G_0'$ . For fp, the finite-strain fits were made separately

<sup>1</sup>Department of Earth and Planetary Materials Science, Graduate School of Science, Tohoku University, Sendai, Miyagi 980-8578, Japan. <sup>2</sup>Japan Synchrotron Radiation Research Institute, Sayo, Hyogo 679-5198, Japan. <sup>3</sup>Department of Earth and Planetary Sciences, Tokyo Institute of Technology, Meguro, Tokyo 152-8551, Japan. <sup>4</sup>Institute for Research on Earth Evolution, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Kanagawa 237-0061, Japan.

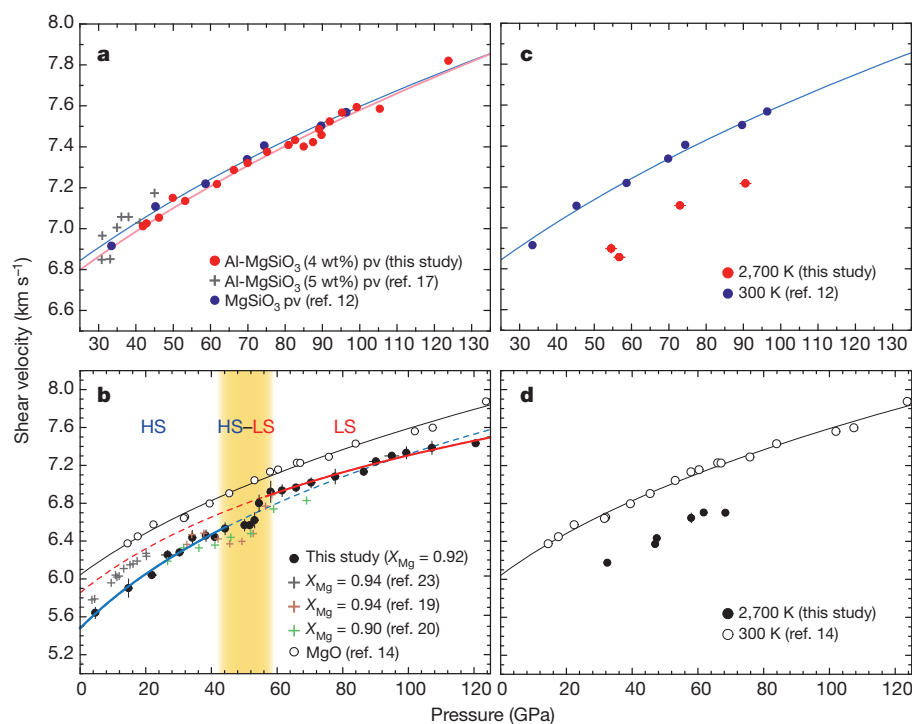


**Figure 1 | Brillouin spectra of lower-mantle phases.** **a**,  $\text{MgSiO}_3$  (+4 wt%  $\text{Al}_2\text{O}_3$ ) pv at 95 GPa and 300 K; **b**,  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  at 86 GPa and 300 K; **c**,  $\text{MgSiO}_3$  pv at 91 GPa and 2,700 K; and **d**,  $\text{MgO}$  at 48 GPa and 2,700 K. TA and

LA indicate transverse and longitudinal acoustic modes of the Brillouin shift, respectively. Ghost peaks are artefacts of the interferometry method used.

for low-pressure (5–40 GPa) and high-pressure (60–121 GPa) because anomalous behaviour was observed at around 50 GPa as a result of spin crossover. The fitting result gives  $G_0 = 113(2)$  GPa and  $G_0' = 2.15(5)$  for the high-spin state, and  $G_0 = 130(2)$  GPa and

$G_0' = 2.04(5)$  for the low-spin state. Extrapolation of the high-spin data to high pressure does not reproduce the low-spin data, supporting the proposal that the spin crossover of iron is not associated with elastic softening<sup>19</sup>.



**Figure 2 | Shear velocity for lower-mantle phases.** **a**,  $\text{MgSiO}_3$  (+4 wt%  $\text{Al}_2\text{O}_3$ ) pv at 300 K. Previous data on  $\text{MgSiO}_3$  (+5 wt%  $\text{Al}_2\text{O}_3$ ) pv<sup>17</sup> and pure  $\text{MgSiO}_3$  pv<sup>12</sup> are also shown. Lines show the regression lines obtained from finite-strain fits. **b**,  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  fp and previous data on fp<sup>19,20,23</sup> and  $\text{MgO}$ <sup>14</sup> at 300 K. The shaded zone indicates the pressure range of the iron spin crossover. Blue and

red lines respectively show the regression lines obtained from finite-strain fit of  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  for high-spin (HS) and low-spin (LS) states of iron. **c**,  $\text{MgSiO}_3$  pv at 300 K (ref. 12) and 2,700 K. **d**,  $\text{MgO}$  at 300 K (ref. 14) and 2,700 K. We note that our data on  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  yield a lower velocity than do those on  $(\text{Mg}_{0.94}\text{Fe}_{0.06})\text{O}$  (ref. 23) owing to higher density. Errors,  $1\sigma$ .

These  $G_0$  and  $G_0'$  values for Al pv and fp are plotted together with previous data<sup>12,17,22–25</sup> in Fig. 3 as a function of iron or aluminium content. Using such relationships, we can estimate  $G_0$  and  $G_0'$  for the representative mantle composition ( $X_{\text{Mg}} = 79 \text{ mol\%}$  for fp and  $X_{\text{Al}_2\text{O}_3} = 4 \text{ wt\%}$  for pv). The high-pressure, high-temperature measurements on pure MgSiO<sub>3</sub> pv and MgO showed that on average the velocity decreased by  $\sim 4\%$  and  $\sim 6\%$ , respectively, in those compounds at 2,700 K, relative to the values at room temperature<sup>12,14</sup> and at pressures corresponding to the lower mantle (Fig. 2c, d). A linear fit of the shear moduli as functions of pressure, combined with the reported values of  $G_0$  and  $G_0'$  (refs 12, 14), provides temperature derivatives of the shear modulus: we find that  $dG/dT = -0.020(1) \text{ GPa K}^{-1}$  for both pv and MgO. The shear strain derivatives of the Grüneisen parameter ( $\gamma$ ) are estimated to be  $\eta_{\text{S0}} = 2.4(2)$  for pv and  $3.0(3)$  for MgO (Supplementary Table 5).

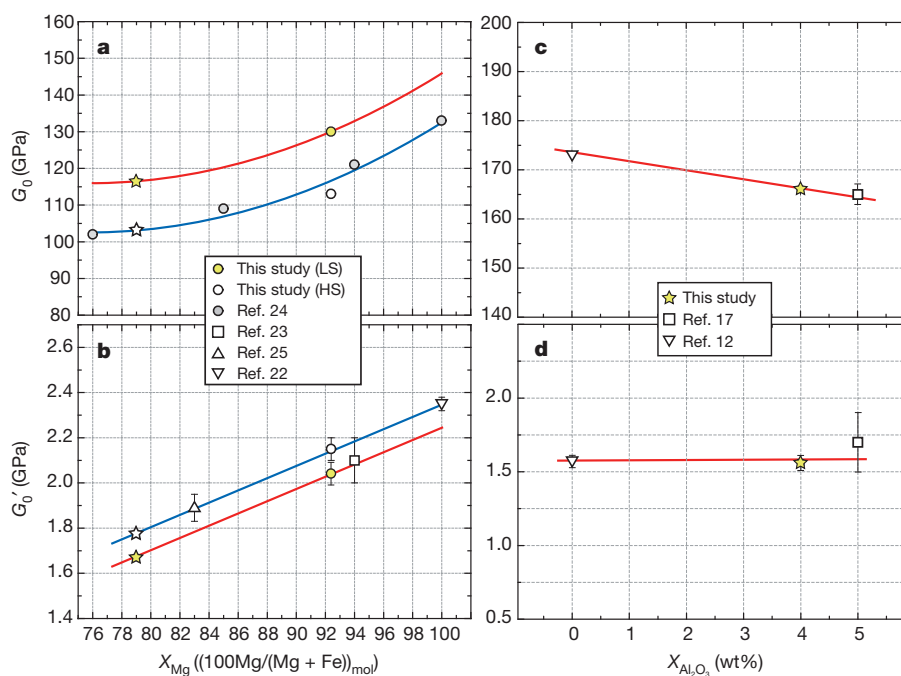
Present measurements performed over a wide  $P$ – $T$  range that covers almost all lower-mantle conditions allow us to constrain the lower-mantle mineralogy<sup>12,14</sup>. We model the lower mantle as a two-phase mixture of pv and fp in a SiO<sub>2</sub>–MgO–FeO–Al<sub>2</sub>O<sub>3</sub> system in which (Al,Fe)-bearing pv contains 4 wt% Al<sub>2</sub>O<sub>3</sub> with  $X_{\text{Mg}} = 94 \text{ mol\%}$  and fp has  $X_{\text{Mg}} = 79 \text{ mol\%}$  (ref. 26). A constant Mg–Fe partitioning coefficient between pv and fp is assumed for the entire lower mantle. The  $V_{\text{S}}$  profiles of pv and fp with the stated compositions were calculated along the typical temperature profiles. Among several different lower-mantle geotherms, here we consider two extreme cases, one for whole-mantle convection<sup>27</sup> and one for layered-mantle convection<sup>28</sup> (Fig. 4a). We did our calculations using the recent compilation of elastic parameters given in the formalism in ref. 21 and using the thermoelastic parameters determined here (Fig. 3 and Supplementary Table 5). The trade-offs between Mg/Fe or Mg/Si and thermal parameters were discussed in previous studies<sup>11,21</sup>. The aggregate  $V_{\text{S}}$  value of the two-phase assemblage was obtained by the Voigt–Reuss–Hill averaging method.

We compared the calculated  $V_{\text{S}}$  profiles with the Preliminary Reference Earth Model<sup>1</sup> (PREM), a one-dimensional global seismic model. For the whole-mantle convection geotherm, the PREM is best fitted by the mixture of 95% pv and 5% fp by volume ( $X_{\text{pv}} = 0.95$ ;

Fig. 4b). The goodness-of-fit for this regression ( $R^2 = 0.996$ ) indicates an excellent agreement with the PREM model. The velocity for fp increases steeply by  $\sim 4\%$  across the spin crossover (Fig. 2b); however, such an anomalous feature is not clear in the calculated profile of the pv + fp mixture, suggesting that the spin crossover of iron in fp may be seismologically unrecognizable. By contrast, the  $V_{\text{S}}$  profile for a peridotitic (pyrolitic) mantle<sup>26</sup> ( $X_{\text{pv}} = 0.80$ ) is lower than the PREM by up to 3.2% throughout the pressure range of the lower mantle, indicating that the conventional peridotitic model is incompatible with the seismological observations.

Such regression suggests that the mantle is not chemically homogeneous and hence implies layered-mantle convection. We therefore applied the layered-mantle convection geotherm and found that the PREM profile is reproduced only by pv without any fp ( $R^2 = 0.995$ ; Fig. 4c), providing further evidence that the mantle is not peridotitic. To verify the internal consistency of those regression results, we also calculated the longitudinal velocity ( $V_{\text{P}} = \sqrt{(K + 4G/3)/\rho}$ ) profile using the bulk modulus ( $K$ ) and density ( $\rho$ ) determined from recent X-ray diffraction data<sup>9,10</sup> ( $G$ , shear modulus). The results show that  $X_{\text{pv}} = 0.93$  ( $R^2 = 0.999$ ; Fig. 4d), consistent with the value estimated from the  $V_{\text{S}}$  profile. CaSiO<sub>3</sub> perovskite (cpv) is also believed to be present in the lower mantle as a host of calcium. There are no experimental data on the sound velocity for cpv under the pressure conditions in the lower mantle, and theoretical predictions are still controversial<sup>29,30</sup>. Nevertheless, cpv is a volumetrically minor ( $\sim 5\%$ ) phase in the lower mantle and the incorporation of cpv thus does not change the present modelling results appreciably.

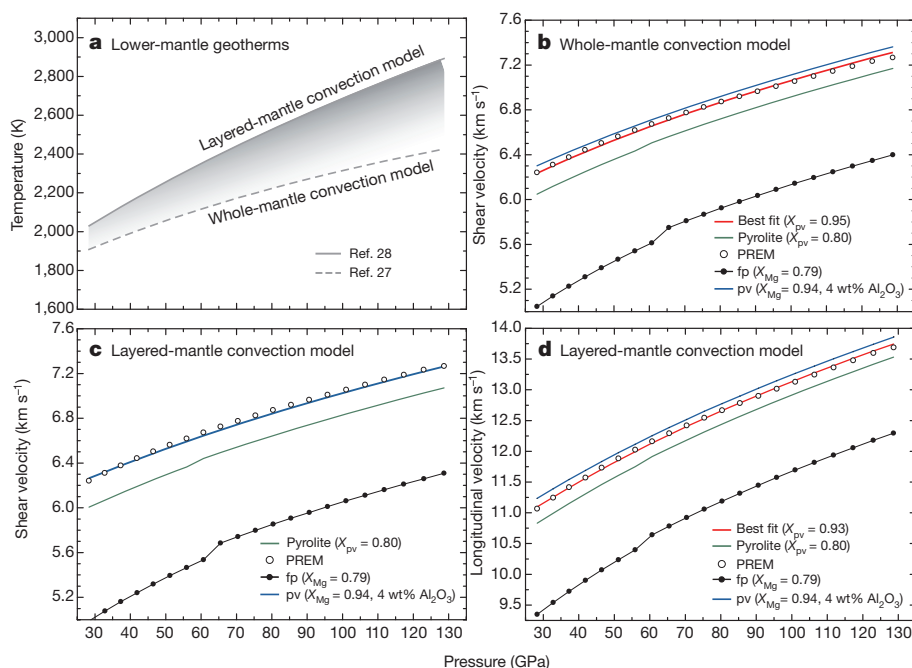
The main uncertainties in our mineralogical modelling reflect the errors in the thermodynamic and thermoelastic parameters we have used. The effect of temperature on shear modulus is most sensitive to  $\eta_{\text{S0}}$  (ref. 21), which was determined for both pv and MgO with an uncertainty of  $\pm 10\%$  in our high-pressure, high-temperature measurements. Also, the Mg–Fe partitioning between pv and fp (quantified by the coefficient  $K_{\text{D}}$ ) remains controversial. Nevertheless, we found that the fitting results for  $X_{\text{pv}}$  are fairly insensitive to these  $\eta_{\text{S0}}$  and  $K_{\text{D}}$  values in modelling reported here. Indeed, the 10% uncertainties in  $\eta_{\text{S0}}$



**Figure 3 | Effect of aluminium and iron on shear modulus and its pressure derivative.** a, b, (Mg,Fe)O fp; and c, d, aluminum-bearing MgSiO<sub>3</sub> pv. Stars represent  $G_0$  (a, c) or  $G_0'$  (b, d) for representative iron and Al<sub>2</sub>O<sub>3</sub> contents of  $X_{\text{Mg}} = 79 \text{ mol\%}$  (ref. 26) for fp and  $X_{\text{Al}_2\text{O}_3} = 4 \text{ wt\%}$  for pv. Lines show the best-

fit trends. Blue and red lines in a and b indicate the trends of (Mg,Fe)O for the high-spin and low-spin states of iron, respectively. The nonlinear relationship between  $G_0$  and iron content in a is based on the results of ref. 24. Errors, 1 $\sigma$ .





**Figure 4 | Lower-mantle geotherms and calculated shear- and longitudinal-wave velocity profiles for whole-mantle and layered-mantle convection models.** **a**, Representative lower-mantle geotherms for whole-mantle convection<sup>27</sup> and layered-mantle convection<sup>28</sup> models. **b–d**, Calculated shear-

(**b**, **c**) and longitudinal-wave (**d**) velocity profiles of fp (black lines) and pv (blue lines) for whole-mantle and layered-mantle convection. The velocity profile for the pyrolitic model is presented as a green line. The best-fit models to the PREM are shown as red curves, indicating the pv-dominant (>93 vol%) lower mantle.

and in  $K_D$  change  $X_{pv}$  by less than 2% and 0.35%, respectively. In addition, the 30% variation in the  $G_0'$  value for fp, which is consistent with almost all previously reported data and corresponds to the uncertainty in the determination here (Fig. 3b), changes  $X_{pv}$  by only ~2%, which is small enough still to support our perovskitic lower-mantle model.

Our results indicate that the conventional peridotitic mantle model is not compatible with the seismic properties of the lower mantle, even considering the experimental uncertainties, and strongly suggest that the lower mantle is dominated by perovskite (>93 vol%) and is therefore silica-rich by comparison with the upper mantle. The lower mantle is chemically distinct, with a near-chondritic Mg/Si ratio (~1.0), which reasonably explains the ‘missing silicon’ problem. The difference in chemical composition between the upper and the lower mantle could be a consequence of fractional crystallization of the magma ocean extending to the deep lower mantle<sup>7</sup> in the early history of Earth. The primordial chemical stratification could be preserved through the subsequent solid-state convection until the present day. Previous mantle convection simulations have demonstrated that a flow pattern evolves from layered- to whole-mantle convection owing to a decrease in the Rayleigh number with cooling<sup>31</sup>. For this reason, layered-mantle convection is presumed to be predominant in the early Earth, indicating that mass transport between the upper and the lower mantle was limited. The seismic tomography images of subducting slabs or upwelling plumes penetrating the boundary between the upper and the lower mantle<sup>32</sup> may not represent whole-mantle convection but rather an intermittent or transitional stage between layered- and whole-mantle convection.

The boundary between the upper and the lower mantle may be of chemical origin. Recent density measurements under relevant  $P$ – $T$  conditions of the lower mantle<sup>10</sup> and computational simulations<sup>11</sup> both also suggest that physical properties of the lower mantle are difficult to reconcile with the peridotitic mantle model but are more consistent with the perovskite-rich lower-mantle models. The lack of internal consistency of the conventionally used pressure scales in the previous high-pressure experiments also requires reconsideration of the interpretation of the 660-km seismic discontinuity. Detailed reassessment

of the post-spinel phase-transition boundary on the basis of a newly established pressure scale has indeed demonstrated that the depth of the 660-km seismic discontinuity does not match the pressure of this phase transition<sup>33</sup>, indicating the incompatibility with an isochemical peridotitic mantle model. The density contrast at the boundary between the peridotitic and the near-chondritic mantle is found to be high enough to allow layered-mantle convection<sup>31</sup>. A dense, stiff, pv-rich lower mantle would promote separate convection and thus inhibit chemical homogenization.

## METHODS SUMMARY

We determined the sound velocities of lower-mantle minerals by using a newly developed system at beamline BL10XU of SPring-8 (ref. 2). The shear-wave velocity data were collected, with reference to the Brillouin scattering measurements, at high pressure in a diamond-anvil cell. A carbon dioxide laser was used to heat the sample, and the sample temperature was determined and monitored by a spectro-radiometric method in high-temperature experiments. Synchrotron X-ray diffraction measurements were simultaneously performed to determine the lattice parameters (volume) of the sample and the pressure standard.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 5 July 2011; accepted 28 February 2012.**

1. Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
2. Murakami, M. *et al.* Development of in-situ Brillouin spectroscopy at high pressure and temperature with synchrotron radiation and infrared laser heating system: application to the Earth's deep interior. *Phys. Earth Planet. Inter.* **174**, 282–291 (2009).
3. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw Hill, 1975).
4. Sun, S. S. Chemical-composition and origin of the Earth's primitive mantle. *Geochim. Cosmochim. Acta* **46**, 179–192 (1982).
5. Ito, E. & Takahashi, E. Postspinel transformations in the system  $Mg_2SiO_4$ – $Fe_2SiO_4$  and some geophysical implications. *J. Geophys. Res.* **94**, 10637–10646 (1989).
6. Allègre, C. J., Poirier, J. P., Humler, E. & Hofmann, A. W. The chemical composition of the earth. *Earth Planet. Sci. Lett.* **134**, 515–526 (1995).
7. Tonks, W. B. & Melosh, H. J. Magma ocean formation due to giant impact. *J. Geophys. Res.* **98**, 5319–5333 (1993).
8. Ringwood, A. E. Significance of the terrestrial Mg/Si ratio. *Earth Planet. Sci. Lett.* **95**, 1–7 (1989).

9. Fei, Y. *et al.* Spin transition and equations of state of (Mg, Fe)O solid solutions. *Geophys. Res. Lett.* **34**, L17307 (2007).
10. Ricolleau, A. *et al.* Density profile of pyrolite under the lower mantle conditions. *Geophys. Res. Lett.* **36**, L06302 (2009).
11. Mattern, E., Matas, J., Ricard, Y. & Bass, J. Lower mantle composition and temperature from mineral physics and thermodynamic modelling. *Geophys. J. Int.* **160**, 973–990 (2005).
12. Murakami, M., Sinogeikin, S. V., Hellwig, H., Bass, J. D. & Li, J. Sound velocity of MgSiO<sub>3</sub> perovskite to Mbar pressure. *Earth Planet. Sci. Lett.* **256**, 47–54 (2007).
13. Murakami, M. *et al.* Sound velocity of MgSiO<sub>3</sub> post-perovskite phase: a constraint on the D'' discontinuity. *Earth Planet. Sci. Lett.* **259**, 18–23 (2007).
14. Murakami, M., Ohishi, Y., Hirao, N. & Hirose, K. Elasticity of MgO to 130 GPa: implications for lower mantle mineralogy. *Earth Planet. Sci. Lett.* **277**, 123–129 (2009).
15. Brodholt, J. P. Pressure-induced changes in the compression mechanism of aluminous perovskite in the Earth's mantle. *Nature* **407**, 620–622 (2000).
16. Jackson, J. M., Zhang, J. & Bass, J. D. Sound velocities and elasticity of aluminous MgSiO<sub>3</sub> perovskite: implications for aluminum heterogeneity in Earth's lower mantle. *Geophys. Res. Lett.* **31**, L10614 (2004).
17. Jackson, J. M., Zhang, J., Shu, J., Sinogeikin, S. V. & Bass, J. D. High-pressure sound velocities and elasticity of aluminous MgSiO<sub>3</sub> perovskite to 45 GPa: implications for lateral heterogeneity in Earth's lower mantle. *Geophys. Res. Lett.* **32**, L21305 (2005).
18. Badro, J. *et al.* Iron partitioning in Earth's mantle: toward a deep lower mantle discontinuity. *Science* **300**, 789–791 (2003).
19. Crowhurst, J. C., Brown, J. M., Goncharov, A. F. & Jacobsen, S. D. Elasticity of (Mg,Fe)O through the spin transition of iron in the lower mantle. *Science* **319**, 451–453 (2008).
20. Marquardt, H., Speziale, S., Reichmann, H. J., Frost, D. J. & Schilling, F. R. Single-crystal elasticity of (Mg<sub>0.9</sub>Fe<sub>0.1</sub>)O to 81 GPa. *Earth Planet. Sci. Lett.* **287**, 345–352 (2009).
21. Stixrude, L. & Lithgow-Bertelloni, C. Thermodynamics of mantle minerals - I. Physical properties. *Geophys. J. Int.* **162**, 610–632 (2005).
22. Kennett, B. L. N. & Jackson, I. Optimal equations of state for mantle minerals from simultaneous non-linear inversion of multiple datasets. *Phys. Earth Planet. Inter.* **176**, 98–108 (2009).
23. Jackson, J. M. *et al.* Single-crystal elasticity and sound velocities of (Mg<sub>0.94</sub>Fe<sub>0.06</sub>)O ferropericlase to 20 GPa. *J. Geophys. Res.* **111**, B09203 (2006).
24. Jacobsen, S. D. *et al.* Structure and elasticity of single-crystal (Mg,Fe)O and a new method of generating shear waves for gigahertz ultrasonic interferometry. *J. Geophys. Res.* **107**, 2037 (2002).
25. Kung, J., Li, B. S., Weidner, D. J., Zhang, J. Z. & Liebermann, R. C. Elasticity of (Mg<sub>0.83</sub>Fe<sub>0.17</sub>)O ferropericlase at high pressure: ultrasonic measurements in conjunction with X-radiation techniques. *Earth Planet. Sci. Lett.* **203**, 557–566 (2002).
26. Jackson, I. & Rigden, S. M. in *The Earth's Mantle: Composition, Structure and Evolution* (ed. Jackson, I.) 405–460 (Cambridge Univ. Press, 1998).
27. Brown, J. M. & Shankland, T. J. Thermodynamic parameters in the Earth as determined from seismic profiles. *Geophys. J. R. Astron. Soc.* **66**, 579–596 (1981).
28. Anderson, O. L. The Earth's core and the phase-diagram of iron. *Phil. Trans. R. Soc. Lond. A* **306**, 21–35 (1982).
29. Li, L. *et al.* Elasticity of CaSiO<sub>3</sub> perovskite at high pressure and high temperature. *Earth Planet. Sci. Lett.* **155**, 249–259 (2006).
30. Stixrude, L., Lithgow-Bertelloni, C., Kiefer, B. & Fumagalli, P. Phase stability and softening in CaSiO<sub>3</sub> perovskite at high pressure. *Phys. Rev. B* **75**, 024108 (2007).
31. Christensen, U. R. & Yuen, D. A. Layered convection induced by phase-transitions. *J. Geophys. Res.* **90**, 10291–10300 (1985).
32. van der Hilst, R., Engdahl, R., Spakman, W. & Nolet, G. Tomographic imaging of subducted lithosphere below northwest Pacific island arcs. *Nature* **353**, 37–43 (1991).
33. Tange, Y., Nishihara, Y. & Tsuchiya, T. Unified analyses for P-V-T equation of state of MgO: a solution for pressure-scale problems in high P-T experiments. *J. Geophys. Res.* **115**, B12203 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We greatly appreciate the comments by I. Jackson. Suggestions from E. Ohtani, C. Bina, S. Karato and S.-H. Shim improved the manuscript. We also thank N. Sata and Y. Asahara for their experimental assistance at SPring-8. This study was performed under the approval of SPring-8 (proposals no. 2008B0099 and 2009A0087).

**Author Contributions** M.M. planned the research and did experimental and analytical work. M.M. and K.H. wrote the paper. All authors were involved in the experiments and discussed the results.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.M. ([motohiko@m.tohoku.ac.jp](mailto:motohiko@m.tohoku.ac.jp)).

## METHODS

**Starting materials.** Gel starting materials were used for experiments on both  $\text{MgSiO}_3$  and aluminous  $\text{MgSiO}_3$  pv. The  $\text{MgSiO}_3$  gel was the same as that used in previous high-pressure experiments on post-pv<sup>34</sup>. The chemical composition of aluminous  $\text{MgSiO}_3$  gel is shown in Supplementary Table 6. For the experiments on  $\text{MgO}$ , fine-grained ( $\sim 1 \mu\text{m}$  in diameter) polycrystalline powder was used<sup>14</sup>. The  $(\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O}$  fp sample was synthesized by mixing stoichiometric amounts of  $\text{MgO}$  and  $\text{Fe}_2\text{O}_3$  powder, heating the mixture overnight at  $1,200^\circ\text{C}$  in reducing conditions using a  $\text{H}_2/\text{CO}_2$  gas-flow furnace, and then quenching. The X-ray diffraction measurement for well-ground fine powder showed a lattice constant of  $a = 4.2209(5) \text{ \AA}$ , indicating that the iron content in this fp sample is  $(\text{Mg}_{0.924}\text{Fe}_{0.076})\text{O}$  (ref. 24), in excellent agreement with the target composition  $((\text{Mg}_{0.92}\text{Fe}_{0.08})\text{O})$ .

**Brillouin scattering measurements.** High-pressure, high-temperature Brillouin scattering measurements of sound velocities were carried out in a laser-heated diamond-anvil cell with a  $60^\circ$  angular aperture. A diode-pumped laser with a wavelength of  $532 \text{ nm}$  was used as a probe beam. An incident laser was focused to a beam size of diameter  $\sim 20 \mu\text{m}$  on the sample. The scattered light was analysed using a six-pass tandem Fabry–Pérot interferometer. All the measurements were performed in a platelet scattering geometry. The scattering angle was calibrated using the glass standard material BK7 in each run. A pre-pressed plate of polycrystalline/gel sample with a thickness of about  $30 \mu\text{m}$  was loaded into a hole ( $100 \mu\text{m}$  in diameter) drilled in a rhenium gasket sandwiched between  $\text{NaCl}$  plates that served as a pressure medium and a thermal insulator.

Polycrystalline pv samples were synthesized *in situ* in a diamond-anvil cell from the gel starting materials by heating with a  $\text{CO}_2$  laser. Pressure was determined on the basis of the Raman  $T_{2g}$  mode of the diamond anvil<sup>35</sup> or the equations of state of  $\text{NaCl}$ <sup>36</sup> at ambient temperature and  $\text{Pt}$ <sup>37</sup> and  $\text{MgO}$ <sup>38</sup> at high temperature. At each pressure, the raw Brillouin spectra of Stokes and anti-Stokes peaks were fitted with a Gaussian peak function to determine the peak positions. We obtained Brillouin spectra on compressed samples in nine, six, three and four separate runs for

aluminous  $\text{MgSiO}_3$  pv, fp,  $\text{MgSiO}_3$  pv and  $\text{MgO}$ , respectively. Each Brillouin spectrum was collected in 2–16 h.

The angle dispersive X-ray diffraction measurements were conducted simultaneously to determine the volume of the sample and the pressure at BL10XU, the synchrotron X-ray source at SPring-8 in the energy range 30–40 keV (Supplementary Fig. 1). The sample temperature was determined and monitored by a spectroradiometric method in the high-temperature experiments. The heating spot of the  $\text{CO}_2$  laser was typically  $40\text{--}50 \mu\text{m}$  in diameter (Supplementary Fig. 2), which was sufficiently larger than the beam spot for the Brillouin measurements. The temporal fluctuation in temperature during heating did not exceed 200 K. As shown in Supplementary Fig. 1, the X-ray diffraction peaks for each sample were very sharp. The two-dimensional X-ray diffraction images showed clear circular Debye rings with a fairly uniform intensity distribution along the circle, indicating that there was no significant grain growth or lattice preferred orientation in the sample at high pressure and high temperature.

**Ambient volume of fp in the low-spin state.** To obtain the zero-pressure shear modulus and its pressure derivative for the fp sample in the low-spin state, we need to determine the volume under ambient conditions ( $V_0$ ). Assuming a linear relationship between  $V_0$  and iron content in low-spin fp as reported in ref. 9, for our sample we estimate that  $V_0 = 74.59 \text{ \AA}^3$ .

34. Murakami, M., Hirose, K., Kawamura, K., Sata, N. & Ohishi, Y. Post-perovskite phase transition in  $\text{MgSiO}_3$ . *Science* **304**, 855–858 (2004).
35. Akahama, Y. & Kawamura, H. High-pressure Raman spectroscopy of diamond anvils to 250 GPa: method for pressure determination in the multimegabar pressure range. *J. Appl. Phys.* **96**, 3748–3751 (2004).
36. Sata, N., Shen, G., Rivers, M. L. & Sutton, S. R. Pressure-volume equation of state of the high-pressure B2 phase of  $\text{NaCl}$ . *Phys. Rev. B* **65**, 104114 (2002).
37. Holmes, N. C., Moriarty, J. A., Gathers, G. R. & Nellis, W. J. The equation of state of platinum to 660 GPa (6.6 Mbar). *J. Appl. Phys.* **66**, 2962–2967 (1989).
38. Dewaele, A., Fiquet, G., Andrault, D. & Hausermann, D. P-V-T equation of state of periclase from synchrotron radiation measurements. *J. Geophys. Res.* **105**, 2869–2877 (2000).



# Evidence of non-random mutation rates suggests an evolutionary risk management strategy

Iñigo Martincorena<sup>1</sup>, Aswin S. N. Seshasayee<sup>1†</sup> & Nicholas M. Luscombe<sup>1,2,3,4</sup>

A central tenet in evolutionary theory is that mutations occur randomly with respect to their value to an organism; selection then governs whether they are fixed in a population. This principle has been challenged by long-standing theoretical models predicting that selection could modulate the rate of mutation itself<sup>1,2</sup>. However, our understanding of how the mutation rate varies between different sites within a genome has been hindered by technical difficulties in measuring it. Here we present a study that overcomes previous limitations by combining phylogenetic and population genetic techniques. Upon comparing 34 *Escherichia coli* genomes, we observe that the neutral mutation rate varies by more than an order of magnitude across 2,659 genes, with mutational hot and cold spots spanning several kilobases. Importantly, the variation is not random: we detect a lower rate in highly expressed genes and in those undergoing stronger purifying selection. Our observations suggest that the mutation rate has been evolutionarily optimized to reduce the risk of deleterious mutations. Current knowledge of factors influencing the mutation rate—including transcription-coupled repair and context-dependent mutagenesis—do not explain these observations, indicating that additional mechanisms must be involved. The findings have important implications for our understanding of evolution and the control of mutations.

The question of whether spontaneous mutations occur randomly has attracted great interest for decades<sup>1–6</sup>. The answer has fundamental implications for evolutionary theory as well as for our understanding of pathogen evolution and certain human diseases. Mutation rates can be modulated by at least three selective forces that can vary in strength depending on genomic location: the cost of deleterious mutations, the need for adaptive mutations, and the cost of fidelity in DNA replication and repair<sup>4,6</sup>. Thus, although mutations are generally assumed to occur independently of their fitness effect, it is conceivable that local mutation rates themselves might evolve, resulting in genomes whose mutations occur non-randomly: more frequently where they are more often advantageous and less frequently where they are most deleterious. Currently, however, there is little evidence that local mutation rates have been optimized during evolution, with the limited exceptions of bacterial contingency loci and somatic hypermutation in the vertebrate immune system.

Our understanding of how the mutation rate varies along a genome has been restricted by the lack of reliable approaches to measure local mutation rates on a large scale. Experimentally, absolute mutation rates can be determined using gene reporters in fluctuation tests, but these are unsuitable for measurements in native genes. Alternatively, in theory, relative mutation rates can be estimated from the accumulation of mutations at selectively neutral positions. Indeed, synonymous or non-coding sites are often used as proxies<sup>7,8</sup> and intriguing correlations between local synonymous substitution rates and gene function or fitness cost have been reported<sup>8,9</sup>. However, because selection can act on these sites through factors like codon-usage preference, RNA-folding

stability and *cis*-regulatory elements, interpretation of these observations in support of optimized mutation rates has remained contentious.

Traditional studies comparing two species suffer from the fundamental limitation that the effects of selection and mutation rate on sequence divergence cannot be distinguished. However, as the two processes leave distinct patterns of polymorphisms, population genetic techniques can be applied to disentangle their relative contributions<sup>10–12</sup>. In examining more than 120,000 single-nucleotide polymorphisms across 34 *E. coli* strains, we exploited this fact to quantify the heterogeneity of the local mutation rate and to test for evidence of evolutionary optimization. Multiple alignments of 2,930 orthologous genes passed stringent phylogenetic filters checking for artefacts like interspecies gene transfer, orthologue misidentification, sequencing errors and misalignments (Supplementary Information, section 2.2). Using these alignments, we then calculated the synonymous diversity of each gene ( $\theta_s = 2N\mu$ ) (Methods and Supplementary Information, section 2.1).

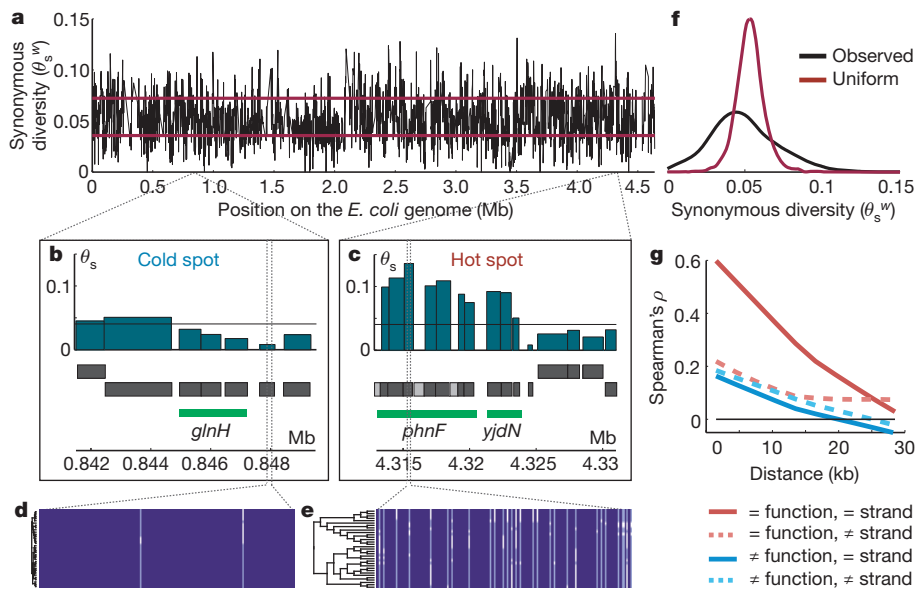
Figure 1a immediately highlights the marked variability in the density of synonymous polymorphisms along the *E. coli* genome. Although the mean synonymous diversity ( $\theta_s$ ) is 0.04, these values vary by more than an order of magnitude from less than 0.002 to more than 0.10 depending on the gene (Fig. 1a–c, f). The observed heterogeneity is much larger than expected under a regime of uniform mutation rate alone (Fig. 1f and Supplementary Information, section 6.1).

The synonymous diversity also shows signs of regional organization (Fig. 1b, c). Mutational hot and cold regions often span entire operons, indicating that gene function has a large effect on  $\theta_s$ . Neighbouring genes tend to have similar  $\theta_s$  values, particularly when they are encoded on the same strand and share common functions (Fig. 1g). A weaker correlation is also apparent for genes with distinct functions and on opposing strands, suggesting that synonymous diversity is affected by regional factors at a resolution of few kilobases.

Although the observed variation in  $\theta_s$  is suggestive, selective and non-selective processes other than mutation rate can also influence synonymous diversity. Synonymous sites in *E. coli* are known to experience purifying selection. The best evidence for this is the negative correlation between inter-species synonymous divergence (dS) and codon usage bias (CUB) ( $R^2 = 0.31$  for *E. coli*–*Salmonella enterica*, Supplementary Fig. 7)<sup>13</sup>. In contrast, within-species diversity in *E. coli* ( $\theta_s$ ) does not correlate with CUB ( $R^2 < 0.01$ , Fig. 2a) demonstrating that, contrary to dS, the variation in  $\theta_s$  is largely neutral with respect to codon usage.

To explain this unexpected result, we applied a mutation-selection-drift model to estimate the strength of selection acting on codon usage<sup>14</sup> (Supplementary Information, section 3.1.1.2). In agreement with previous reports<sup>15</sup>, we find that this is relatively weak for most genes ( $\gamma_{\text{mean}} = -0.41$ ,  $\gamma_{\text{std}} = 0.40$ , Fig. 2b). These estimated selection coefficients are too small to affect within-species  $\theta_s$  substantially, but large enough to affect inter-species dS<sup>14</sup> (Fig. 2a and Supplementary Fig. 8). The other prominent source of selection at synonymous sites,

<sup>1</sup>EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK. <sup>2</sup>Okinawa Institute of Science & Technology, 1919-1 Tancha, Onna-son, Kunigami-gun, Okinawa 904-0495, Japan. <sup>3</sup>UCL Genetics Institute, Department of Genetics, Environment and Evolution, University College London, Gower Street, London WC1E 6BT, UK. <sup>4</sup>Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK. <sup>†</sup>Present address: National Centre for Biological Sciences, TIFR, GKVK, Bellary Road, Bangalore 560065, India.



**Figure 1 | Synonymous diversity along the *E. coli* genome is heterogeneous.**

**a**, Distribution of the 2,930  $\theta_s$  values (using  $\theta_s^W$ , see Supplementary Information, section 6.1) plotted along the *E. coli* K12 MG1655 genome showing that synonymous diversity is highly variable. The red lines indicate the upper and lower 2.5% limits of the expected distribution under a regime of uniform mutation rate. Forty-one per cent of  $\theta_s^W$  values fall outside of the expected range. **b, c**, Details of two mutational hot and cold regions displaying the genomic coordinates, annotated genes and their respective  $\theta_s$  values. Each gene is shown as a black box on the sense or anti-sense strands (grey for genes with no  $\theta_s$  available). Green horizontal lines indicate operons, labelled by the name of their lead gene. **d, e**, Multiple-sequence alignments of 180-base-pair

segments of two genes, **(d)** *dps* ( $\theta_s = 0.008$ ;  $TD_{syn} = 0.35$ ; RNA-Seq level = 139.9), and **(e)** *phnL* ( $\theta_s = 0.136$ ;  $TD_{syn} = 0.33$ ; RNA-Seq level = 1.4). **f**, Probability density distributions of  $\theta_s$  values: observed (black) and expected under a uniform mutation rate alone (red). **g**, Fitted lines showing the decay of the correlation between pairs of  $\theta_s$  values with their genomic distance. Neighbouring genes tend to have similar  $\theta_s$ , particularly if they share similar functions and they are encoded on the same strand (red line). Weaker but significant correlations are apparent even for functionally unrelated genes (blue) and those on opposing strands (dotted lines). An analogous pattern is observed for  $\theta_s'$  (Supplementary Fig. 22).

mRNA-folding stability in the 5' end of genes<sup>16,17</sup>, does not explain the variation of  $\theta_s$  either (Supplementary Information, section 3.1.2).

In addition to direct selection, positive and purifying selection at linked sites (known as hitchhiking and background selection, respectively) affect local neutral diversity in eukaryotes<sup>18</sup>. However, these processes have no effect on synonymous diversity here. Hitchhiking would at most result in lower  $\theta_s$  in regions under positive selection, which is inconsistent with our observations (Supplementary Information, section 3.1.4). Furthermore, in bacteria where recombination occurs without crossovers and instead resembles gene conversion<sup>19</sup>, partial linkage is conserved throughout the entire genome (Supplementary Figs 11 and 16) and background selection cannot preferentially reduce local  $\theta_s$  values (Fig. 2c and Supplementary Information, section 3.1.4).

The results so far indicate that, contrary to dS, the genome-wide variation of  $\theta_s$  is largely neutral with respect to the strongest known selective forces at synonymous sites in bacteria and that much stronger forces are needed to bias  $\theta_s$  significantly. Nevertheless, to account for any possible source of selection, including unknown ones, we applied neutrality tests based on the site-frequency spectrum (Tajima's *D*, and Fu and Li's *D* and *F*), which provide relative measures of selection that are independent of the local mutation rate (Fig. 2d)<sup>19</sup>. Using realistic forward simulations of genomes undergoing non-crossover recombination and variable amounts of selection, we studied the impact of selection on  $\theta_s$  and the allele spectrum. The tests show that Tajima's *D* would detect any traces of selection strong enough to bias  $\theta$  (Supplementary Information, section 3.1.3).

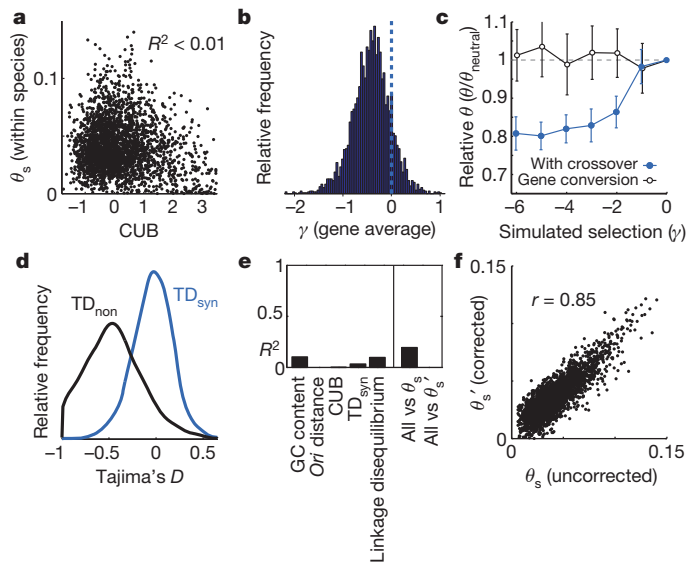
Regression on Tajima's *D* at synonymous sites ( $TD_{syn}$ ) indicates that selection from all sources explains less than 5% of the variance in  $\theta_s$  (Fig. 2e), even accounting for the noise in estimating  $TD_{syn}$  using few polymorphisms (Supplementary Fig. 9). This is in agreement with the independently obtained observations above that CUB and mRNA-folding stability have minimal impact on the observed variation of  $\theta_s$ .

We also explored the influence of non-selective factors on  $\theta_s$ , including GC content and within-species homologous recombination. All potential biases together only account for 20% of the variation in  $\theta_s$  (Fig. 2e), suggesting that there is a large underlying variation in the neutral mutation rate along the *E. coli* genome.

Despite the weak association between the above factors and  $\theta_s$ , we eliminated these small biases using locally weighted scatterplot smoothing (LOWESS) regression for 2,659 genes (Fig. 2f, Supplementary Fig. 19 and Supplementary Information, section 3.4) to avoid ambiguity in the interpretation of the functional associations below. We challenged the regression approach using a cross-validation test (Fig. 3e, top panel, and Supplementary Information, section 3.1.3.6) and extensive forward simulations (Supplementary Information, section 3.1.3.8). They showed that selection bias on  $\theta_s$  can be comprehensively removed by regression on  $TD_{syn}$  even in evolutionary regimes with strong linkage disequilibrium, gene conversion, non-random sampling of bacterial strains and stronger selective forces than those observed here (Supplementary Information, section 3.1.3, and Supplementary Figs 12–14). Thus, the genome-wide variation of the corrected  $\theta_s'$  values can be considered effectively neutral.

To assess whether the genome-wide variation in mutation rate shows signs of evolutionary optimization, we studied the association between  $\theta_s'$  and the strength of purifying selection on the protein sequences ( $TD_{non}$ ). Figure 3a provides initial evidence of the association between the neutral mutation rate and selection: namely, that proteins under stronger purifying selection experience a lower rate.

We then studied the relationship between the mutation rate and functional genomic data for *E. coli* K12. Genes with lower  $\theta_s'$  tend to be essential for survival in rich media<sup>20</sup> compared with those of higher  $\theta_s'$  (Fig. 3b). Furthermore, mutationally cold genes generally encode for vital cellular functions such as processes related to central energy metabolism and the respiratory chain (Supplementary Table 1). Mutationally hot genes are associated with metabolic pathways expressed at lower



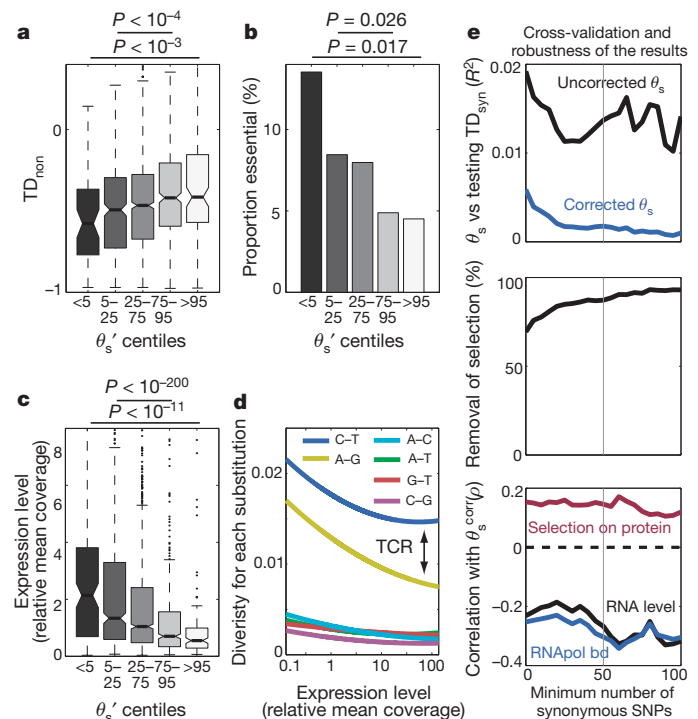
**Figure 2 | Selective and non-selective factors have only small effects on the variation of  $\theta_s$ .** **a**, Within-species  $\theta_s$  does not display the traditional strong correlation with codon usage bias. Only a small fraction of genes with extremely high CUB display a certain decrease in  $\theta_s$  from selection on codon usage. **b**, Histogram of the weighted average selection coefficient for each gene. These values of  $\gamma$  are too small to cause significant deviations of  $\theta_s$  (see also Supplementary Fig. 8). **c**, Effect of background selection on the relative local neutral diversity in forward simulations. Background selection cannot affect the local variation of neutral diversity in the presence of non-crossover recombination. Circles represent median values and error bars show the 95% confidence interval of the medians. **d**, Probability density distribution of normalized Tajima's  $D$  at synonymous ( $TD_{syn}$ ; blue) and at non-synonymous sites ( $TD_{non}$ ; black) across all genes. The distributions indicate strong purifying selection at non-synonymous sites relative to synonymous sites. **e**, Bar plots displaying the proportion ( $R^2$ ) of the variance of  $\theta_s$  explained by five selective and non-selective potential biasing factors. All factors combined explain around 20% of the variance in  $\theta_s$ . Their effect is eliminated from  $\theta_s'$  after LOWESS correction. **f**, Scatter plot showing the correlation between  $\theta_s$  and  $\theta_s'$ . Removal of the potential biases has only a moderate effect on the overall variation of  $\theta_s$ . Analogous results are obtained with or without the adjustment.

levels or used less frequently, such as amino-acid biosynthesis and catabolism of specific compounds. Interestingly, hot genes do not include antigens and other genes expected to experience frequent positive selection; in fact, we were unable to detect any association between  $\theta_s'$  and the density of sites under positive selection (Supplementary Information, section 6.4, and Supplementary Fig. 23). Thus, together our observations suggest that purifying, rather than positive, selection has driven the evolution of the mutation rate along the core *E. coli* genome.

Next, we studied the relationship between mutation rate and gene expression using transcriptomic (RNA-Seq) and RNA-polymerase-binding data (Supplementary Information, section 2.4). There is a clear negative correlation between  $\theta_s'$  and transcription levels (Fig. 3c), with the coldest genes (bottom 5%  $\theta_s'$ ) showing nearly fourfold higher transcription than the hottest (top 5%  $\theta_s'$ ). Earlier studies have reported associations between dS and expression levels, but they were generally accompanied by correlations with CUB reflecting the action of selection<sup>13</sup>. Other studies accounting for CUB also suggested that the mutation rate is lower at highly expressed genes<sup>21</sup>, but the roles of additional sources of selection remained unresolved. Our data now demonstrate this in the absence of confounding biases.

Finally, because functional properties are often related to expression levels, we performed a multiple regression analysis. This confirmed that the mutation rate independently associates with expression levels, selection on protein sequence and gene function (Supplementary Information, section 6.7).

Given that the process of transcription is known to be mutagenic<sup>22–24</sup>, the negative association between expression and mutation rate is unexpected. This implies the presence of compensatory mechanisms that preferentially protect or repair highly expressed loci in *E. coli*. Indeed, we do detect the action of the transcription-coupled repair pathway<sup>25</sup> (Fig. 3d). However, this mechanism alone is insufficient, because the non-transcribed strand also displays a dependence between expression and mutation rate (Fig. 3d), and molecular experiments have reported that transcription-induced mutagenesis occurs in the presence of transcription-coupled repair. Therefore additional mechanisms that generally target highly expressed genes, but are not directly coupled with the transcriptional machinery, must exist.



**Figure 3 | Variation in the mutation rate shows functional dependence.** **a–c**, Box- and bar-plots displaying the relationship between gene function and mutation rate. Genes were classified as displaying from very low (less than the fifth centile) to very high (greater than the 95th centile)  $\theta_s'$  values.  $P$  values correspond to **(a, c)** Wilcoxon rank-sum test and **(b)** Fisher's exact test. **a**, Genes with higher mutation rate tend to experience weaker purifying selection at the protein level (measured using  $TD_{non}$  for genes with at least five non-synonymous single nucleotide polymorphisms). **b**, Genes with lower mutation rate show greater tendency to be essential for survival in rich media. **c**, Genes with lower mutation rate are generally more highly expressed (using RNA-Seq data from *E. coli* K12, Supplementary Information, section 2.4). The central mark of each box-plot represents the median, the edges of the box are the 25th and 75th centiles, and the notches are the 95% confidence interval of the median. Whiskers extend to the most extreme data points within the range. Dots show outliers. **d**, Fitted lines showing the relationship between  $\theta_s'$  per substitution type and expression level. The direction of the substitution is not shown as we used a reversible evolutionary model. Because the C-to-T transition (G-to-A in the opposite strand) is the most common mutation, the gap between the C–T (blue) and G–A (yellow) lines indicates the strand asymmetry caused by the action of transcription-coupled repair (TCR) on the transcribed strand. **e**, Lines showing the results of cross-validation tests challenging the correction for selection bias (Supplementary Information, section 3.1.3.6). Top, the correlation of  $\theta_s$  and  $TD_{syn}$  before and after adjustment by an independent set of  $TD_{syn}$  values shows that selection bias is removed. Middle, proportion of selection bias eliminated ( $R^2_{after}/R^2_{before}$ ), given different minimum numbers of synonymous single nucleotide polymorphisms in the training set. Bottom, strength of functional associations of  $\theta_s$  values at different levels of correction, showing that results are unaffected by the adjustment. SNPs, single nucleotide polymorphisms.



Our observations suggest that purifying selection has driven the evolution of the local point mutation rate in *E. coli* to reduce the risk of deleterious mutations. This contrasts with most earlier theoretical work that proposed variants of bet-hedging in which frequent positive selection in changing environments leads to the emergence of hypermutators<sup>1–3</sup>. Instead our observations are in line with an evolutionary risk-management strategy<sup>26</sup> in which sustained stronger purifying selection at specific genes favours individuals with preferential protection or repair at these loci, even at a cost of reduced protection of other genes (see Supplementary Information, section 4.3, for a detailed description of the model). In this way, the rate of deleterious mutations in the genome can be efficiently reduced without excessive investment in protection or repair. In addition, this could increase the rate of non-deleterious mutations, so raising the adaptive potential of the population in case of an environmental change.

We can only speculate about the molecular mechanisms underlying the localized reduction in spontaneous mutations. DNA-binding proteins and DNA repair pathways are obvious candidates, especially as there is increasing evidence for locus-dependent control of the latter<sup>27,28</sup>. In eukaryotes, the best-known sources for mutation rate heterogeneity are sequence dependent; however, analysis of sequence-enrichments around our data set of synonymous polymorphisms provided little indication of context-dependent mutagenesis (Supplementary Fig. 24).

Surprising negative correlations between expression and the numbers of substitutions were recently reported in several human cancers, for which both mutational<sup>29</sup> and selective<sup>30</sup> interpretations have been proposed. Understanding the extent and mechanisms of the modulation of the local mutation rate in different organisms will be important for research into evolution, human diseases, mutagenesis and repair.

## METHODS SUMMARY

A total of 3,420 one-to-one orthologues present in at least 26 (75%) of 34 *E. coli* strains were identified by a strict reciprocal best BLAST-hit approach. They were aligned using PRANK-F with a codon substitution model. Multiple phylogenetic filters were applied to discard alignments potentially affected by artefacts. Extensive quality controls were performed on the remaining 2,930 alignments. A  $\theta_s$  value was calculated for each alignment using a codon model (OmegaMap).

CUB was calculated using the codon adaptation index or the fraction of optimal codons. The scaled selection coefficients ( $\gamma = 2Ns$ ) of selection on codon usage were estimated as in ref. 14. To minimize the impact of selection on mRNA folding, all analyses were repeated after trimming the ends of alignments yielding analogous results. Tajima's  $D$  was used to estimate selection from any source, and the sensitivity of the approach was tested by multiple independent tests including forward simulations (Supplementary Information, section 3.1.3). The variation of  $\theta_s$  explained by selective and non-selective biases was removed in 2,659 genes using nonlinear LOWESS regression. This set was used in all functional analyses.

The corrections were challenged using cross-validation tests and extensive forward simulations. The cross-validation showed that  $TD_{syn}$  biases can be removed to near completion and that increased removal of selection bias does not affect the functional results (Fig. 3e). SFS\_CODE was used to simulate multiple populations of genomes comprising 11 loci under different levels of purifying and diversifying selection, and subject to realistic levels of mutation and non-crossover recombination. For background selection and hitchhiking we simulated coding sequences with selection acting only at non-synonymous sites, and quantified its impact on neutral diversity at synonymous sites under crossover recombination or gene conversion. Homologous recombination was estimated independently by linkage disequilibrium and phylogenetic consistency, yielding analogous results.

Received 27 October 2011; accepted 29 February 2012.

Published online 22 April 2012.

- Kimura, M. On the evolutionary adjustment of spontaneous mutation rates. *Genet. Res.* **9**, 23–24 (1967).
- Levins, R. Theory of fitness in a heterogeneous environment. VI. The adaptive significance of mutation. *Genetics* **56**, 163–178 (1967).
- Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**, 24–33 (1994).

- Sniegowski, P. D., Gerrish, P. J., Johnson, T. & Shaver, A. The evolution of mutation rates: separating causes from consequences. *BioEssays* **22**, 1057–1066 (2000).
- Tenaillon, O., Taddei, F., Radman, M. & Matic, I. Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res. Microbiol.* **152**, 11–16 (2001).
- Pal, C., Macia, M. D., Oliver, A., Schachar, I. & Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**, 1079–1081 (2007).
- Hodgkinson, A., Ladoukakis, E. & Eyre-Walker, A. Cryptic variation in the human mutation rate. *PLoS Biol.* **7**, e1000027 (2009).
- McVean, G. T. & Hurst, L. D. Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. *Nature* **386**, 388–392 (1997).
- Chuang, J. H. & Li, H. Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLoS Biol.* **2**, E29 (2004).
- Braverman, J. M., Hudson, R. R., Kaplan, N. L., Langley, C. H. & Stephan, W. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics* **140**, 783–796 (1995).
- O'Fallon, B. D. A method to correct for the effects of purifying selection on genealogical inference. *Mol. Biol. Evol.* **27**, 2406–2416 (2010).
- Bustamante, C. D., Nielsen, R. & Hartl, D. L. Maximum likelihood and Bayesian methods for estimating the distribution of selective effects among classes of mutations using DNA polymorphism data. *Theor. Popul. Biol.* **63**, 91–103 (2003).
- Drummond, D. A. & Wilke, C. O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* **134**, 341–352 (2008).
- McVean, G. A. & Charlesworth, B. A. A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genet. Res.* **74**, 145–158 (1999).
- Hartl, D. L., Moriyama, E. N. & Sawyer, S. A. Selection intensity for codon bias. *Genetics* **138**, 227–234 (1994).
- Eyre-Walker, A. & Bulmer, M. Reduced synonymous substitution rate at the start of enterobacterial genes. *Nucleic Acids Res.* **21**, 4599–4603 (1993).
- Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
- Andolfatto, P. Adaptive hitchhiking effects on genome variability. *Curr. Opin. Genet. Dev.* **11**, 635–641 (2001).
- Touchon, M. *et al.* Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* **5**, e1000344 (2009).
- Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Systems Biol.* **2**, 0008, doi:10.1038/msb4100050 (2006).
- Eyre-Walker, A. & Bulmer, M. Synonymous substitution rates in enterobacteria. *Genetics* **140**, 1407–1412 (1995).
- Ochman, H. Neutral mutations and neutral substitutions in bacterial genomes. *Mol. Biol. Evol.* **20**, 2091–2096 (2003).
- Beletskii, A. & Bhagwat, A. S. Transcription-induced mutations: increase in C to T mutations in the nontranscribed strand during transcription in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **93**, 13919–13924 (1996).
- Klapacz, J. & Bhagwat, A. S. Transcription-dependent increase in multiple classes of base substitution mutations in *Escherichia coli*. *J. Bacteriol.* **184**, 6866–6872 (2002).
- Francino, M. P., Chao, L., Riley, M. A. & Ochman, H. Asymmetries generated by transcription-coupled repair in enterobacterial genes. *Science* **272**, 107–109 (1996).
- Wagner, A. Risk management in biological evolution. *J. Theor. Biol.* **225**, 45–57 (2003).
- Tu, Y., Tornaletti, S. & Pfeifer, G. P. DNA repair domains within a human gene: selective repair of sequences near the transcription initiation site. *EMBO J.* **15**, 675–683 (1996).
- Hoeghe, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. & Jentsch, S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**, 135–141 (2002).
- Pleasant, E. D. *et al.* A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**, 191–196 (2010).
- Lee, W. *et al.* The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* **465**, 473–477 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Ackermann, S. Brenner, G. Dougan, A. Eyre-Walker, N. Goldman, B. Lenhard, J. Marioni, J. Parkhill, O. Tenaillon, C. Tyler-Smith and F. Uhlmann for their suggestions during the preparation of this manuscript. The work was funded by EMBL, the Spanish Ministry of Science and Innovation and the Caja Madrid Foundation.

**Author Contributions** I.M. and N.M.L. conceived the study; I.M. designed and performed the analyses; A.S.N.S. and N.M.L. provided advice; I.M. and N.M.L. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to I.M. ([martino@ebi.ac.uk](mailto:martino@ebi.ac.uk)) or N.M.L. ([luscombe@ebi.ac.uk](mailto:luscombe@ebi.ac.uk)).

# Restoration of vision after transplantation of photoreceptors

R. A. Pearson<sup>1</sup>, A. C. Barber<sup>1</sup>, M. Rizzi<sup>1</sup>, C. Hippert<sup>1</sup>, T. Xue<sup>3</sup>, E. L. West<sup>1</sup>, Y. Duran<sup>1</sup>, A. J. Smith<sup>1</sup>, J. Z. Chuang<sup>4</sup>, S. A. Azam<sup>1</sup>, U. F. O. Luhmann<sup>1</sup>, A. Benucci<sup>2</sup>, C. H. Sung<sup>4</sup>, J. W. Bainbridge<sup>1</sup>, M. Carandini<sup>2</sup>, K.-W. Yau<sup>3</sup>, J. C. Sowden<sup>5</sup> & R. R. Ali<sup>1,6</sup>

Cell transplantation is a potential strategy for treating blindness caused by the loss of photoreceptors. Although transplanted rod-precursor cells are able to migrate into the adult retina and differentiate to acquire the specialized morphological features of mature photoreceptor cells<sup>1</sup>, the fundamental question remains whether transplantation of photoreceptor cells can actually improve vision. Here we provide evidence of functional rod-mediated vision after photoreceptor transplantation in adult *Gnat1*<sup>-/-</sup> mice, which lack rod function and are a model of congenital stationary night blindness<sup>2</sup>. We show that transplanted rod precursors form classic triad synaptic connections with second-order bipolar and horizontal cells in the recipient retina. The newly integrated photoreceptor cells are light-responsive with dim-flash kinetics similar to adult wild-type photoreceptors. By using intrinsic imaging under scotopic conditions we demonstrate that visual signals generated by transplanted rods are projected to higher visual areas, including V1. Moreover, these cells are capable of driving optokinetic head tracking and visually guided behaviour in the *Gnat1*<sup>-/-</sup> mouse under scotopic conditions. Together, these results demonstrate the feasibility of photoreceptor transplantation as a therapeutic strategy for restoring vision after retinal degeneration.

So far there have been no convincing reports of photoreceptor-cell transplantation actually improving the recipient's vision. This may be due to the relatively low numbers of new rod photoreceptors successfully transplanted in previous studies (typically fewer than 1,000 cells)<sup>1,3–6</sup>. To establish that new rods truly can improve vision, we optimized the rod-transplantation procedure to increase the number of newly integrated photoreceptor cells in wild-type mice. The donor-cell population was rod-photoreceptor precursors identified by their expression of green fluorescent protein (GFP) under control of the promoter for the rod-specific transcription factor, *Nrl*. We obtained maximum integration together with optimal recipient retinal histology after the transplantation of 200,000 fluorescence-activated cell sorted *Nrl-GFP*<sup>+</sup> rod precursors, taken from postnatal day (P) 4–8 *Nrl-GFP* donor mice, by subretinal injection to both the superior and the inferior retina. This resulted in a 20- to 30-fold increase in the number of integrated rod photoreceptors compared with previous studies, with up to 26,000 new rods within the outer nuclear layer of recipient adult wild-type mice (16,759 ± 1,705 cells; Fig. 1a and Supplementary Fig. 1). Integrated cells were found predominantly around the injection sites, but were distributed over more than 50% of the retinal area. Thus, up to 16% of donor cells integrated into the host outer nuclear layer (see Supplementary Information for estimation of integrated cell number). Supplementary Fig. 1 summarizes how this was achieved.

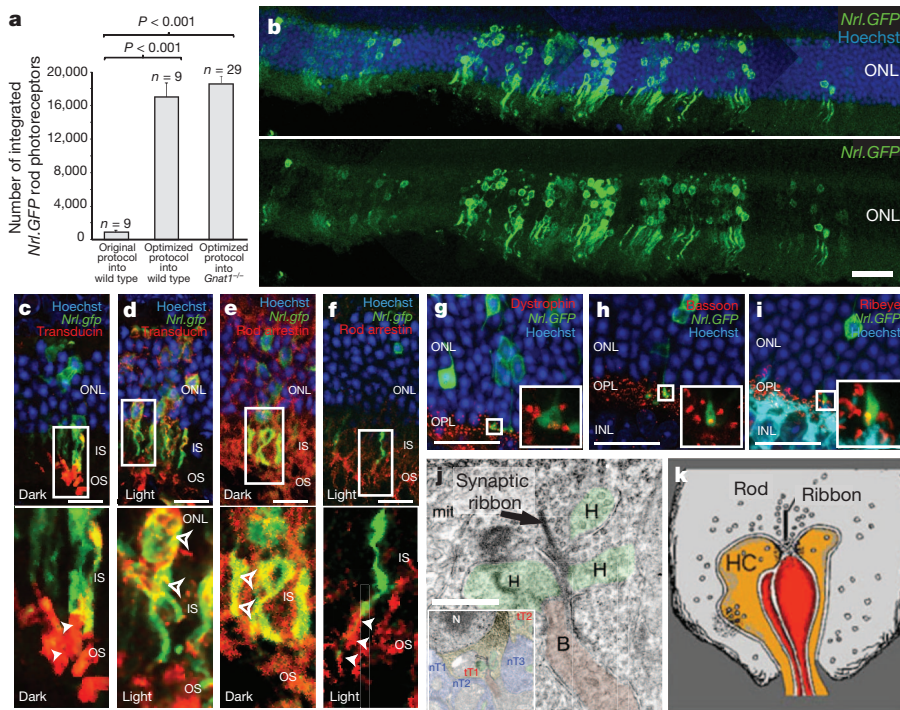
To test the functionality of the transplanted rod photoreceptors, we selected a murine model in which improvement of rod vision could be assessed definitively; the *Gnat1*<sup>-/-</sup> mouse lacks rod  $\alpha$ -transducin (*Gnat1*), a protein essential for rod phototransduction. It has no rod

function<sup>2</sup> or behavioural responses to scotopic visual stimuli; it displays, however, near-normal cone histology<sup>8</sup>, function and behavioural responses to photopic visual stimuli (Supplementary Fig. 2 and refs 2, 9) and shows no loss of cones with time<sup>8</sup>. *Nrl-GFP*<sup>+</sup> rod precursors integrated into the adult *Gnat1*<sup>-/-</sup> retina in numbers very similar to those observed in wild-type recipients (18,300 ± 1474; maximum 32,015 integrated cells; Fig. 1a). Integrated cells were appropriately located within the outer nuclear layer and morphologically very similar to wild-type rods (Fig. 1b–i and Supplementary Fig. 3), correctly expressing rod  $\alpha$ -transducin (Fig. 1c, d), which is absent in endogenous *Gnat1*<sup>-/-</sup> rods, together with all other mature rod markers examined (rod arrestin (Fig. 1e, f), rhodopsin, recoverin, phosducin (data not shown)). They demonstrated outer segment formation (Fig. 1b–f) and appropriate light-dependent translocation of rod  $\alpha$ -transducin (Fig. 1c, d) and arrestin (Fig. 1e, f)<sup>10</sup> between the outer segments and the cell body/synapse. Most (>80%) of *Nrl-GFP*<sup>+</sup> rods located within the outer nuclear layer displayed synaptic boutons. All those examined by immunohistochemistry (15–20 cells per marker) appropriately expressed the rod ribbon synapse proteins ribeye, bassoon and/or dystrophin<sup>11</sup> (Fig. 1g–i). Finally, ultrastructural analysis confirmed that integrated *Nrl-GFP*<sup>+</sup> rods typically formed the classic triad synapse with endogenous horizontal and bipolar neurons (Fig. 1j, k; synapses observed for 51/55 cells examined; see Supplementary Information).

To test for light responses in *Nrl-GFP*<sup>+</sup> rods correctly integrated within *Gnat1*<sup>-/-</sup> recipients, we performed suction-pipette recordings<sup>12</sup> (Fig. 2a, b). Pigment bleached by the GFP-excitation light used during *Nrl-GFP*<sup>+</sup> cell searching was regenerated with 9-*cis*-retinal before recording (see Supplementary Information). To control for the effects of the bleaching/regeneration procedure on the speed of the dim-flash response kinetics and the sensitivity of wild-type rods (Fig. 2b), we also recorded from non-injected wild-type retinas with and without the excitation-light/pigment-regeneration procedure (Fig. 2a, b). All recorded *Nrl-GFP*<sup>+</sup> rods with an intact inner segment, cell body and synaptic terminal showed light responses ( $n = 9/9$ ; Fig. 2a, b). Their saturated photocurrent (4.6 ± 0.5 pA) was similar to that of wild-type rods with or without bleaching/regeneration treatment (Fig. 2b; 6.4 ± 0.7 and 6.1 ± 0.8 pA, respectively). Dim-flash response kinetics, indicated by time-to-peak ( $t_{peak}$ ) and integration time ( $t_i$ ), were similar between *Nrl-GFP*<sup>+</sup> and wild-type rods subjected to bleaching/regeneration (see Fig. 2b). Individual *Nrl-GFP*<sup>+</sup> rods showed variations in sensitivity, with one cell being as sensitive as bleached/regenerated wild-type rods and the other three over a log unit less sensitive (Fig. 2b). This variation may be real, or simply reflect the small sample size (owing to technical challenges in obtaining this parameter in bleached/regenerated *Nrl-GFP*<sup>+</sup> rods). For each *Nrl-GFP*<sup>+</sup> rod recorded, several adjacent GFP<sup>-ve</sup> rods were also tested; none gave detectable responses (Fig. 2c), thus ruling out light signals originating from cones in the *Gnat1*<sup>-/-</sup> retina.

<sup>1</sup>Department of Genetics UCL Institute of Ophthalmology, University College London, 11–43 Bath Street, London, EC1V 9EL, UK. <sup>2</sup>Department of Visual Neuroscience, UCL Institute of Ophthalmology, University College London, 11–43 Bath Street, London, EC1V 9EL, UK. <sup>3</sup>Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. <sup>4</sup>Dyson Vision Research Institute, Department of Ophthalmology, Department of Cell and Developmental Biology, Weill Medical College of Cornell University, New York, New York 10021, USA. <sup>5</sup>Developmental Biology Unit, UCL Institute of Child Health, University College London, 30 Guilford Street, London, WC1N 1EH, UK. <sup>6</sup>Molecular Immunology Unit, UCL Institute of Child Health, University College London, 30 Guilford Street, London, WC1N 1EH, UK.

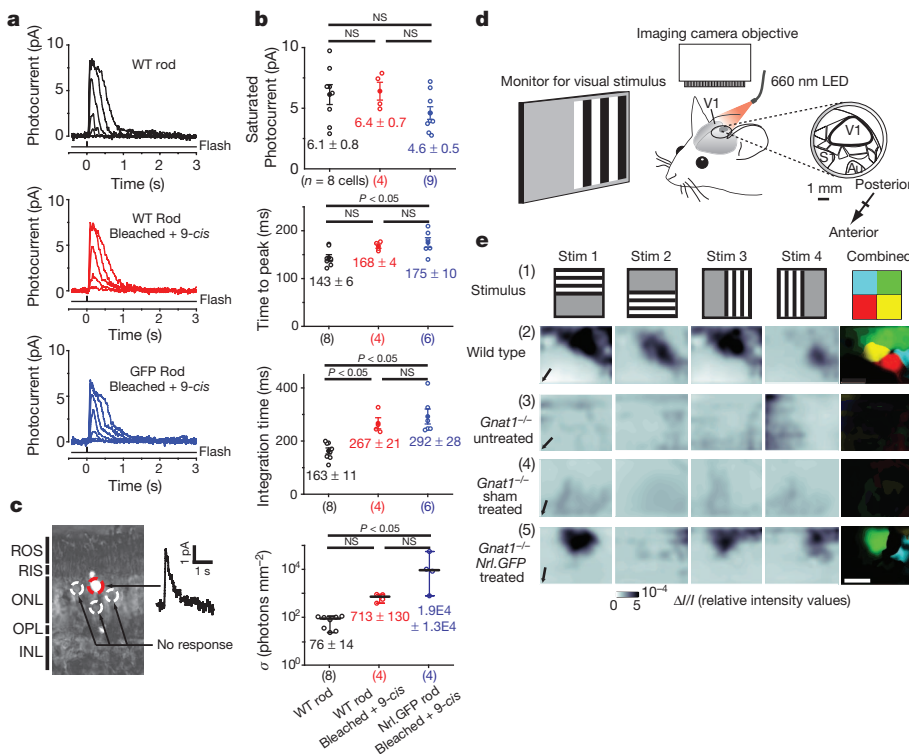




**Figure 1 | Improved transplantation protocols significantly improve photoreceptor integration into the adult *Gnat1*<sup>-/-</sup> model of retinal dysfunction.** **a**, *Nrl-GFP*<sup>+</sup> rod-photoreceptor integration using new and previously published protocols<sup>1</sup>; mean  $\pm$  s.e.m., analysis of variance (ANOVA);  $n$ , number of eyes. **b**, Typical example of integrated *Nrl-GFP*<sup>+</sup> rods (green). Scale bar, 50  $\mu$ m. **c-f**, Integrated *Nrl-GFP*<sup>+</sup> rods expressed rod- $\alpha$ -transducin (c, d; red) and rod-arrestin (e, f; red) and demonstrated correct, counter-directional light-mediated translocation of these proteins (bottom panels in c-f, respectively). **g-i**, Integrated *Nrl-GFP*<sup>+</sup> rods formed spherule synapses and expressed dystrophin (g), bassoon (h) and ribeye (i) (all red). Inserts, 3  $\mu$ m projections of regions marked. IS, inner segments; OS, outer segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer. Scale bar, 10  $\mu$ m. **j**, Electron micrographs of low- (inset) and high-power views (consecutive sections), showing *Nrl-GFP*<sup>+</sup> rod terminals (tT1, tT2). tT1 formed classic triad with horizontal axon terminals (H, orange) and bipolar dendritic terminal (B, red). Endogenous rod terminals (nT; blue) were DAB negative. Arrows, synaptic ribbons; scale bars, 500 nm. **k**, Schematic of rod-triad synapse, reproduced with permission from Webvision (H. Kolb *et al.*, <http://webvision.med.utah.edu/>).

Electroretinography (ERG) provides a gross measure of light-mediated trans-retinal function, averaged across the whole retina. *Gnat1*<sup>-/-</sup> mice received dual transplants of *Nrl-GFP*<sup>+</sup> rod precursors into one eye and equivalent sham injections (age-matched *Gnat1*<sup>-/-</sup> cells or vehicle) or no injection into the contralateral eye (Supplementary Fig. 4). ERGs were recorded weekly 3–6 weeks after transplantation using a double-masked protocol (see Supplementary Information), before the eyes were assessed for *Nrl-GFP*<sup>+</sup> rod-photoreceptor integration. Despite robust integration (7,135–26,616 integrated cells per eye;  $n=6$ ), ERG responses were not detected upon scotopic stimulation of procedure *Gnat1*<sup>-/-</sup> animals (Supplementary Fig. 4a). In contrast,

such responses were readily recordable in wild-type animals (Supplementary Fig. 4f). These results were not unexpected because control experiments with an adeno-associated (AAV2/8) viral vector to deliver a *Gnat1* transgene showed that a scotopic ERG response was recordable only after transduction of approximately 150,000 photoreceptor cells, but not of 60,000 rods (Supplementary Fig. 4d–g). Robust responses were observed in all groups after photopic stimulation with no significant differences between sham and *Nrl-GFP*<sup>+</sup>-treated eyes (Supplementary Fig. 4h). Thus, proper functioning of all retinæ was maintained throughout and was not impaired by the transplantation procedures.



**Figure 2 | Transplanted *Nrl-GFP*<sup>+</sup> rod photoreceptors are light-responsive and project light information to the visual cortex.** **a**, Flash-response families from single dark-adapted wild type (black), bleached/regenerated wild type (red) and bleached/regenerated-*Nrl-GFP*<sup>+</sup> (blue) rods. See Supplementary Information for flash intensities. **b**, Comparison of saturated response amplitude and dim-flash response parameters. Individual data points are shown (open circles) with mean  $\pm$  s.e.m. (ANOVA, top three panels) or median  $\pm$  range (Kruskal–Wallis, bottom panel);  $n$ , number of cells;  $\sigma$ , half-saturating flash intensity. Not all parameters were obtainable for each cell. **c**, Representative retinal slice showing light-sensitive *Nrl-GFP*<sup>+</sup> rod (red circle) and surrounding non-responsive GFP negative rods (white circles). **d**, Schematic of optical intrinsic imaging set-up (see Supplementary Information). Au, auditory cortex; S1, somatosensory cortex. **e**, Visual stimuli (Stim) (1) elicited optical signals in V1 only from *Nrl-GFP*<sup>+</sup>-treated *Gnat1*<sup>-/-</sup> (5) and wild type (2), but not untreated (3) or sham-injected (4), *Gnat1*<sup>-/-</sup> eyes. Far right, overlapping parts of the four stimuli were colour-coded (1). Only responses present for two overlapping stimuli were considered genuine sensory-evoked signals and represented with the corresponding colour (2–5). Scale bar, 1 mm.

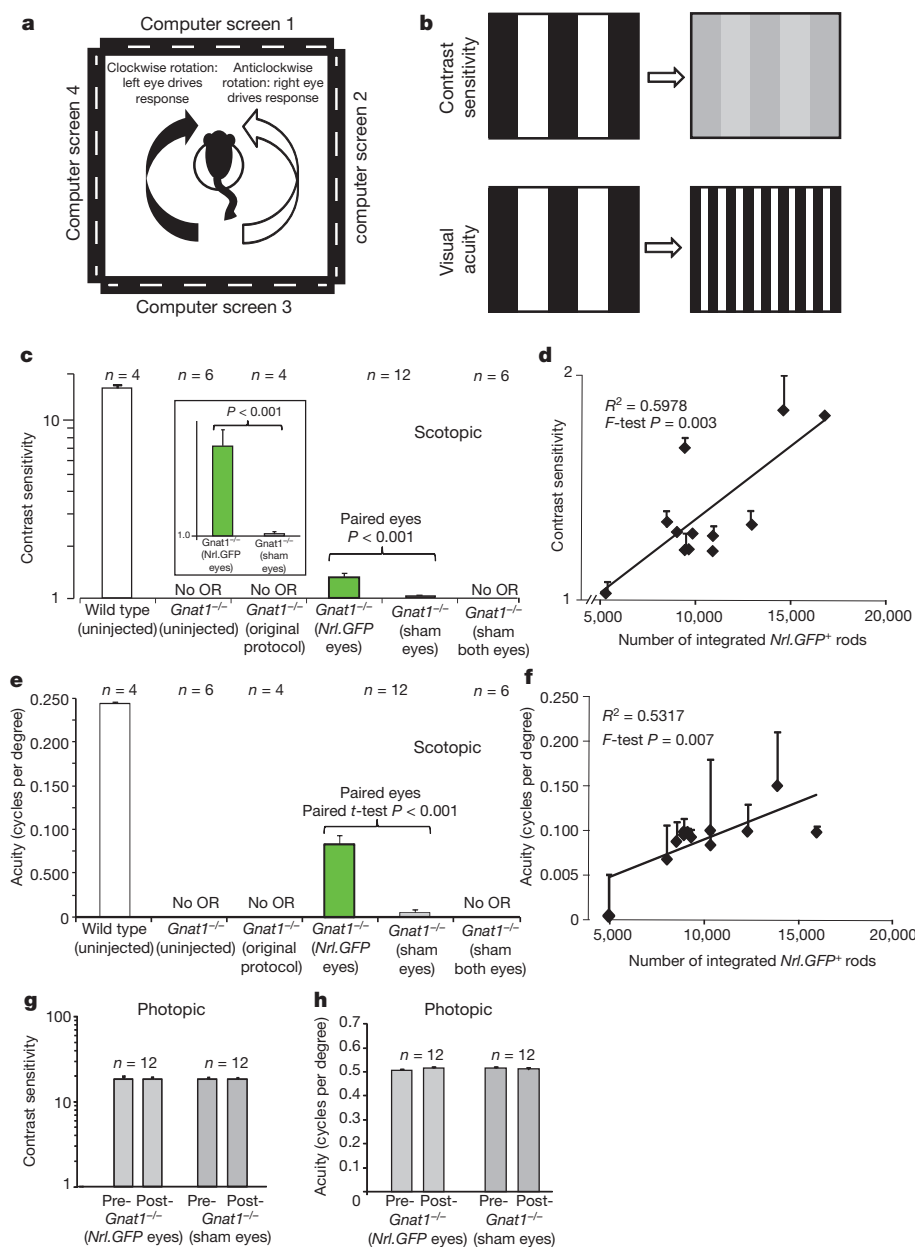


Given the magnification of retinal signals shown by the visual pathway, we reasoned that photoresponses recorded from transplanted *Nrl-GFP*<sup>+</sup> rods, although too small in amplitude or too few in number to be detected by ERG, could be transmitted to the brain and generate activity in defined areas of the visual cortex. We therefore performed optical intrinsic imaging<sup>13,14</sup> of primary visual cortex (Fig. 2d, e and Supplementary Fig. 2b). Four overlapping stimuli covering most of the visual field of the stimulated eye were used and consisted of flickering black/white bars on a grey background (Fig. 2e, row i). When presented to wild-type mice ( $n = 5$ ), such stimuli always evoked strong signals in well-defined areas of the visual cortex under both scotopic (Fig. 2e, row ii) and photopic (Supplementary Fig. 2b) conditions. Normal retinotopy was observed, as indicated by the different cortical locations of the responses to each stimulus (colour maps; Fig. 2e and Supplementary Fig. 2b). In contrast, no detectable response was observed after scotopic stimulation of untreated *Gnat1*<sup>-/-</sup> mice ( $n = 7$ ) (Fig. 2e, row iii, and Supplementary Figs 2b and 5). We assessed the impact of transplantation by giving *Gnat1*<sup>-/-</sup> animals *Nrl-GFP*<sup>+</sup> rod precursors or age-matched *Gnat1*<sup>-/-</sup> (sham) cells, as before. We observed no cortical activity in response to scotopic visual stimuli

presented to sham-injected eyes ( $n = 3$ ; Fig. 2e, row iv, and Supplementary Fig. 5), demonstrating that neither injection procedure nor presence of non-functional cells led to false signals in the neocortex. However, presentation of the same stimuli to eyes receiving *Nrl-GFP* transplants led to robust responses in all animals ( $n = 5$ ; Fisher's exact test  $P = 0.0175$ ) (Fig. 2e, row v, and Supplementary Fig. 5); partly overlapping stimuli gave responses in likewise partly overlapping areas of the neocortex, with some preservation of retinotopy in the areas of integration (colour maps; Fig. 2e, row v, and Supplementary Fig. 5). Robust integration (10,782–31,075 cells) was observed in all animals displaying cortical activity. Because transplanted cells predominantly integrate near the injection site, it is reasonable that only some stimuli led to a reliable cortical response (compare stimuli 1, 3, 4 with stimulus 2, Fig. 2e, row v). Whenever possible, after using scotopic stimuli we assessed photopic stimuli; such stimuli elicited clear responses in all groups.

To address the paramount question of whether photoreceptor transplantation confers improvements in vision, we assessed visual function first by measuring optomotor head-tracking responses to a rotating grating (Fig. 3a, b; see Supplementary Information)<sup>9,15,16</sup>.

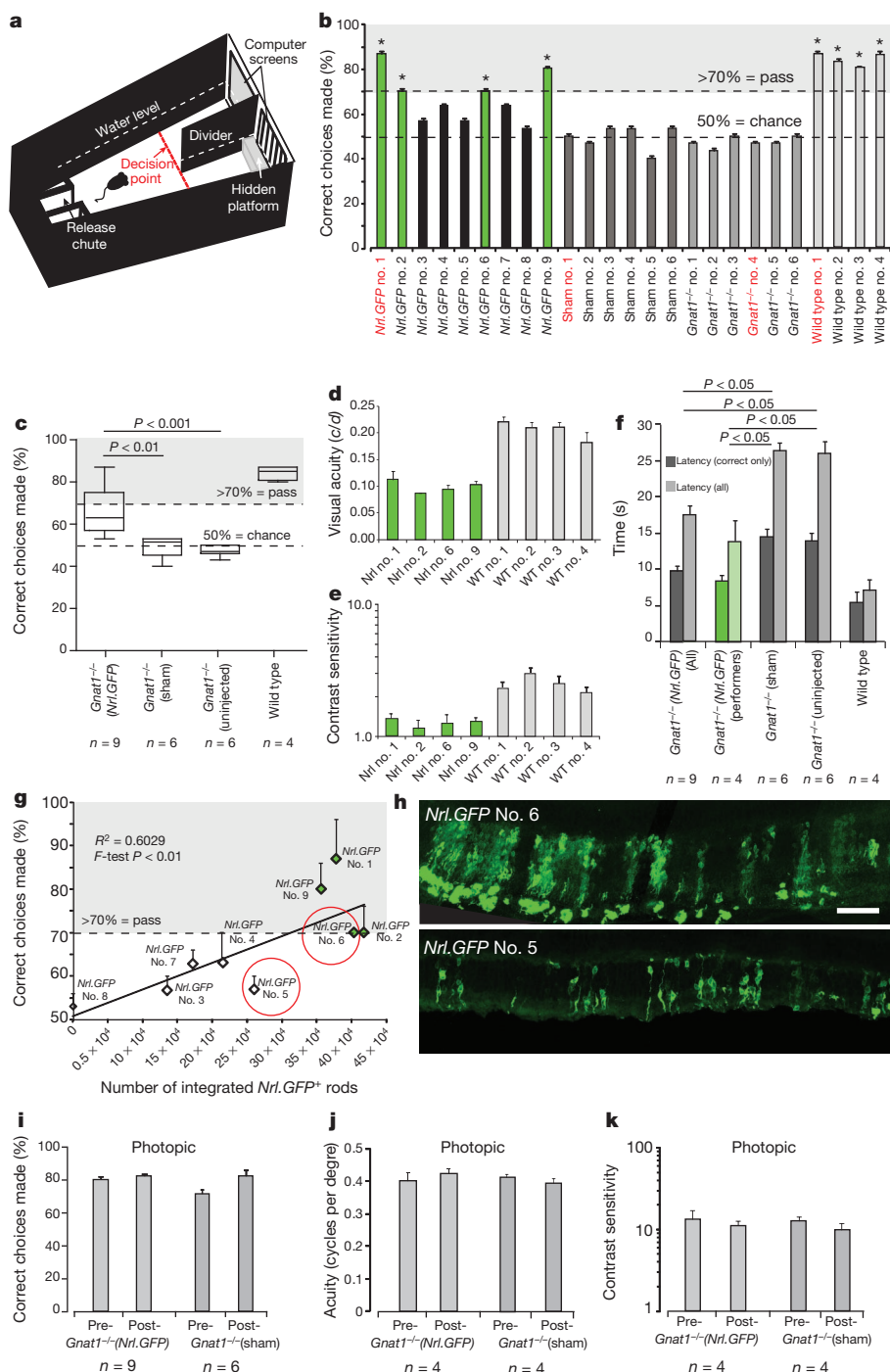
**Figure 3 | Rescue of scotopic optokinetic head-tracking behaviour in *Nrl-GFP*-treated *Gnat1*<sup>-/-</sup> mice.** **a**, The Optomotry<sup>TM</sup> set-up (see Supplementary Information). **b**, Measures of visual function include contrast sensitivity and visual acuity. **c, d**, Scotopic contrast sensitivity and visual acuity threshold measurements for *Nrl-GFP* (green bars) or sham- (*Gnat1*<sup>-/-</sup>) treated *Gnat1*<sup>-/-</sup> eyes, and the averages of left and right eyes for *Gnat1*<sup>-/-</sup> mice receiving original protocol transplants<sup>1</sup> or sham injections to both eyes, or untreated *Gnat1*<sup>-/-</sup> or wild type (white bars) controls. OR, optomotor response. Paired *t*-test. **e, f**, Scatter plots of contrast sensitivity and visual acuity against integrated *Nrl-GFP*<sup>+</sup> rod number. **g, h**, Photopic contrast sensitivity and visual acuity for *Nrl-GFP*<sup>-</sup> (light grey) and sham-treated (dark grey) eyes before and after transplantation. Means  $\pm$  s.e.m.;  $n$ , number of animals.



*Gnat1*<sup>-/-</sup> mice received dual transplants of *Nrl-GFP*<sup>+</sup> rod precursors into one eye and age-matched *Gnat1*<sup>-/-</sup> cells into the contralateral eye, as previously described. Additional cohorts received (1) the same, but using our original protocol<sup>1</sup>, (2) dual *Gnat1*<sup>-/-</sup> sham injections to both eyes or (3) no injection. As expected, no head-tracking behaviour was observed in any *Gnat1*<sup>-/-</sup> mouse at scotopic luminance levels (Fig. 3c, e) and normal head tracking occurred under photopic conditions (Supplementary Fig. 2c, d) before transplantation. After transplantation, scotopic optomotor responses were only seen in *Nrl-GFP*-treated eyes; mean contrast-sensitivity threshold for *Nrl-GFP*-treated animals was  $1.3 \pm 0.1$  compared with  $15.0 \pm 0.6$  in wild type (Fig. 3c). Murine visual acuity is poorer under scotopic than photopic conditions ( $0.245 \pm 0.01$  versus  $0.488 \pm 0.003$  cycles per degree in wild type; see also ref. 9) but a visual-acuity threshold of

$0.097 \pm 0.007$  cycles per degree was recorded in *Nrl-GFP*-treated *Gnat1*<sup>-/-</sup> mice under the same conditions (Fig. 3e). Moreover, there was a significant positive correlation between both contrast sensitivity (*F*-test,  $P < 0.01$ ; Fig. 3d) and visual acuity (*F*-test,  $P < 0.05$ ; Fig. 3f) and the number of integrated *Nrl-GFP*<sup>+</sup> rods. Photopic optomotor responses, assessed after scotopic testing, remained unchanged (Fig. 3g, h).

Finally, we assessed vision using the visually guided water-maze task test<sup>17</sup>, which requires cognitive processing of visual information to associate a grating with escape from a Y-shaped water maze. Under photopic conditions, *Gnat1*<sup>-/-</sup> mice learned and completed the task as well as wild type ( $\geq 70\%$  correct responses; Fig. 4a, Supplementary Fig. 2e–g and Supplementary Information) but performed no better than chance under scotopic conditions. *Gnat1*<sup>-/-</sup> mice then received



**Figure 4** | *Nrl-GFP*-treated *Gnat1*<sup>-/-</sup> mice can solve the visually guided water-maze task under scotopic conditions. **a**, Schematic of water-maze apparatus (adapted from ref. 22; see Supplementary Information). Mice were trained to associate striped grating with escape from water by a hidden platform. An animal 'passes' a trial by crossing the red line (decision point) on the side of the divider with the striped grating. **b**, Pass rate of *Nrl-GFP*-treated (black), sham-injected (dark grey) and non-injected (mid grey) *Gnat1*<sup>-/-</sup> and non-injected wild-type (light grey) mice. *Nrl-GFP*-treated animals with a pass-rate of at least 70% are shown in green throughout. Mouse numbers in red refer to mice shown in Supplementary Movie. **c**, Average performance rate of all groups. **d**, Visual acuity and **e**, contrast sensitivity measurements for responders from *Nrl-GFP*-treated (green) and wild-type (light grey) groups. **f**, Swim-time latencies (time-to-platform) for all (light grey) and correct choice-only (dark grey) trials. **g**, Ability to solve water-maze task plotted against integrated *Nrl-GFP* photoreceptor number. **h**, Examples of integration in animals that successfully (top; *Nrl-GFP*-treated, number 6) or unsuccessfully (bottom; *Nrl-GFP*-treated, number 5) solved the task, as indicated in **g** (circled, red). Scale bar, 100  $\mu$ m. **i–k**, Pass rate (**i**), visual acuity (**j**) and contrast sensitivity (**k**) for *Nrl-GFP*-treated (light grey bars) and sham-injected (dark grey bars) *Gnat1*<sup>-/-</sup> mice before and after transplantation under photopic conditions. Means  $\pm$  s.e.m.; ANOVA; *n*, number of animals.

(1) dual injections of *Nrl-GFP*<sup>+</sup> rod precursors to both eyes, (2) sham injections to both eyes or (3) remained untreated. After transplantation, four out of nine *Gnat1*<sup>-/-</sup> animals receiving *Nrl-GFP*<sup>+</sup> rod precursors correctly completed the task in at least 70% of trials under scotopic conditions (green bars, Fig. 4b; Supplementary Movie). In contrast, all sham and untreated *Gnat1*<sup>-/-</sup> mice performed no better than chance (50%). As a group, *Nrl-GFP*-treated animals were significantly better at solving the water-maze task than either control group (Fig. 4c). The four *Nrl-GFP*-treated *Gnat1*<sup>-/-</sup> animals able to solve the task correctly further had a contrast sensitivity of  $1.3 \pm 0.04$  and visual acuity of  $0.09 \pm 0.005$  cycles per degree (compared with contrast sensitivity of  $3.0 \pm 0.24$  and visual acuity of  $0.206 \pm 0.005$  cycles per degree, respectively, for wild-type controls; Fig. 4d, e). These same animals solved the task with an average swim time of  $15.1 \pm 0.8$  s (correct choice swim-time latency  $9.6 \pm 0.8$  s), compared with  $26.4 \pm 1.1$ ,  $26.0 \pm 1.2$  and  $7.1 \pm 1.4$  s for sham-injected *Gnat1*<sup>-/-</sup>, untreated *Gnat1*<sup>-/-</sup> and wild-type mice, respectively (Fig. 4f). There was a significant positive correlation between ability to solve the task and integrated *Nrl-GFP*<sup>+</sup> rod number (*F*-test,  $P < 0.01$ ; Fig. 4g, h). Cell integration in those animals that performed best in this task was often clustered, rather than widely distributed across the retina, suggesting that proximity of integrated cells to one another may be important, in addition to absolute number. For example, *Nrl-GFP*-treated animals numbers 1 and 2 had similar numbers of integrated rod photoreceptors but number 1 had qualitatively more closely grouped clusters of cells (data not shown). Ability to perform the water-maze task under photopic conditions, assessed after scotopic testing, remained unchanged in all groups after transplantation (Fig. 4i–k).

Although the scotopic visual function recorded here in transplanted *Gnat1*<sup>-/-</sup> mice is lower than in wild-type animals, such sensitivity is still impressive given that integrated photoreceptors account for less than 1% of the total rods in the retina. However, other studies have demonstrated a remarkable sensitivity from very few light-sensitive cells<sup>18</sup>, whereas clinical studies have shown useful vision despite there being no detectable ERG response from the retina<sup>19</sup>. Many additional steps are required before these findings can be translated to the clinic, but the results presented here demonstrate for the first time that transplanted rod-photoreceptor precursors can integrate into a dysfunctional adult retina and, by directly connecting with the host retinal circuitry, truly improve vision.

## METHODS SUMMARY

The transplantation procedures were performed as described previously<sup>1,3–5</sup> with modifications, as described in Supplementary Fig. 1 and Supplementary Information. All recipients were adult (6–8 weeks) at time of transplantation and functional assessments were made 4–6 weeks after transplantation. Cellular, retinal and behavioural assessments of visual function were assessed using suction-pipette recordings<sup>12,20</sup>, ERGs<sup>21</sup>, optical imaging<sup>13,14</sup>, optomotor<sup>9,15,16</sup> and watermaze<sup>17</sup> testing and previously published protocols with amendments as detailed in Supplementary Information.

Received 22 June 2011; accepted 28 February 2012.

Published online 18 April 2012.

- MacLaren, R. E. *et al.* Retinal repair by transplantation of photoreceptor precursors. *Nature* **444**, 203–207 (2006).
- Calvert, P. D. *et al.* Phototransduction in transgenic mice after targeted deletion of the rod transducin  $\alpha$ -subunit. *Proc. Natl Acad. Sci. USA* **97**, 13913–13918 (2000).
- West, E. L. *et al.* Pharmacological disruption of the outer limiting membrane leads to increased retinal integration of transplanted photoreceptor precursors. *Exp. Eye Res.* **86**, 601–611 (2008).
- Pearson, R. A. *et al.* Targeted disruption of outer limiting membrane junctional proteins (Crb1 and ZO-1) increases integration of transplanted photoreceptor precursors into the adult wild-type and degenerating retina. *Cell Transplant.* **19**, 487–503 (2010).

- West, E. L. *et al.* Long-term survival of photoreceptors transplanted into the adult murine neural retina requires immune modulation. *Stem Cells* **28**, 1997–2007 (2010).
- Bartsch, U. *et al.* Retinal cells integrate into the outer nuclear layer and differentiate into mature photoreceptors after subretinal transplantation into adult mice. *Exp. Eye Res.* **86**, 691–700 (2008).
- Akimoto, M. *et al.* Targeting of GFP to newborn rods by *Nrl* promoter and temporal expression profiling of flow-sorted photoreceptors. *Proc. Natl Acad. Sci. USA* **103**, 3890–3895 (2006).
- Maeda, T. *et al.* A critical role of CaBP4 in the cone synapse. *Invest. Ophthalmol. Vis. Sci.* **46**, 4320–4327 (2005).
- Umino, Y., Solessio, E. & Barlow, R. B. Speed, spatial, and temporal tuning of rod and cone vision in mouse. *J. Neurosci.* **28**, 189–198 (2008).
- Elias, R. V. *et al.* Temporal kinetics of the light/dark translocation and compartmentation of arrestin and  $\alpha$ -transducin in mouse photoreceptor cells. *Mol. Vis.* **10**, 672–681 (2004).
- Garner, C. C., Kindler, S. & Gundelfinger, E. D. Molecular determinants of presynaptic active zones. *Curr. Opin. Neurobiol.* **10**, 321–327 (2000).
- Nikonov, S. S. *et al.* Physiological features of the S- and M-cone photoreceptors of wild-type mice from single-cell recordings. *J. Gen. Physiol.* **127**, 359–374 (2006).
- Grinvald, A. *et al.* Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature* **324**, 361–364 (1986).
- Schuetz, S., Bonhoeffer, T. & Hubener, M. Mapping retinotopic structure in mouse visual cortex with optical imaging. *J. Neurosci.* **22**, 6549–6559 (2002).
- Alexander, J. J. *et al.* Restoration of cone vision in a mouse model of achromatopsia. *Nature Med.* **13**, 685–687 (2007).
- Prusky, G. T. *et al.* Rapid quantification of adult and developing mouse spatial vision using a virtual optomotor system. *Invest. Ophthalmol. Vis. Sci.* **45**, 4611–4616 (2004).
- Prusky, G. T., West, P. W. & Douglas, R. M. Behavioral assessment of visual acuity in mice and rats. *Vision Res.* **40**, 2201–2209 (2000).
- Lagali, P. S. *et al.* Light-activated channels targeted to ON bipolar cells restore visual function in retinal degeneration. *Nature Neurosci.* **11**, 667–675 (2008).
- Berson, E. L. Long-term visual prognoses in patients with retinitis pigmentosa: the Ludwig von Sallmann lecture. *Exp. Eye Res.* **85**, 7–14 (2007).
- Fu, Y. *et al.* Quantal noise from human red cone pigment. *Nature Neurosci.* **11**, 565–571 (2008).
- Tan, M. H. *et al.* Gene therapy for retinitis pigmentosa and Leber congenital amaurosis caused by defects in AIP1: effective rescue of mouse models of partial and complete Aip1 deficiency using AAV2/2 and AAV2/8 vectors. *Hum. Mol. Genet.* **18**, 2099–2114 (2009).
- Wong, A. A. & Brown, R. E. Age-related changes in visual acuity, learning and memory in C57BL/6J and DBA/2J mice. *Neurobiol. Aging* **28**, 1577–1593 (2007).

Supplementary Information is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by the Medical Research Council UK (G03000341), the Wellcome Trust (082217), the Royal Society (RG080398), the British Retinitis Pigmentosa Society (GR566) and The Miller's Trust. R.A.P. is a Royal Society University Research Fellow. R.R.A. is partly funded by the Department of Health's National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital and Alcon Research Institute. J.C.S. is supported by Great Ormond Street Hospital Children's Charity. T.X. and K.-W.Y. were supported by a US National Institutes of Health grant (EY06837) and the António Champalimaud Vision Award (Portugal). C.H.S. was supported by grants EY11307, EY016805, Research to Prevent Blindness. M.C. holds the GlaxoSmithKline/Fight for Sight Chair in Visual Neuroscience and is supported by the European Research Council. We thank Y. Umino and the late R. Barlow for advice and training with optomotor recordings, A. Eddoudi for FACS assistance, G. Holder for advice on ERG recordings, C. Hogg for light calibrations, L. Cao for technical suggestions on suction-pipette recordings, and the Department of Genetics, UCL Institute of Ophthalmology, for discussions on the data.

**Author Contributions** R.A.P. contributed to the concept, design, execution and analysis of all experiments, funding, and wrote the manuscript. A.C.B. contributed to the design, execution and analysis of experiments except single-cell recordings, electron microscopy and behavioural assessments. M.R., A.B. and M.C. contributed to the design, execution and analysis of the intrinsic optical-imaging experiments. C.H., A.J.S. and S.A.A. contributed to the design and execution of the AAV gene supplementation studies, and C.H. the execution of the water-maze tests. Y.D. contributed to the histological processing and execution of the water-maze tests. J.W.B., E.L.W. and U.F.O.L. contributed to the optimization of transplantation protocols. T.X. performed the single-cell recordings. K.W.Y. contributed to the design and analysis of the single-cell recordings. C.H.S. and J.Z.C. performed the ultrastructural analysis. J.C.S. and R.R.A. contributed to the concept and design of the experiments, funding and to manuscript writing.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.A.P. ([rachael.pearson@ucl.ac.uk](mailto:rachael.pearson@ucl.ac.uk)) or R.R.A. ([r.ali@ucl.ac.uk](mailto:r.ali@ucl.ac.uk)).



# A unifying model for mTORC1-mediated regulation of mRNA translation

Carson C. Thoreen<sup>1,2,3</sup>, Lynne Chantranupong<sup>3,4,5</sup>, Heather R. Keys<sup>3,4,5</sup>, Tim Wang<sup>3,4</sup>, Nathanael S. Gray<sup>1,2</sup> & David M. Sabatini<sup>3,4,5</sup>

The mTOR complex 1 (mTORC1) kinase nucleates a pathway that promotes cell growth and proliferation and is the target of rapamycin, a drug with many clinical uses<sup>1</sup>. mTORC1 regulates messenger RNA translation, but the overall translational program is poorly defined and no unifying model exists to explain how mTORC1 differentially controls the translation of specific mRNAs. Here we use high-resolution transcriptome-scale ribosome profiling to monitor translation in mouse cells acutely treated with the mTOR inhibitor Torin 1, which, unlike rapamycin, fully inhibits mTORC1 (ref. 2). Our data reveal a surprisingly simple model of the mRNA features and mechanisms that confer mTORC1-dependent translation control. The subset of mRNAs that are specifically regulated by mTORC1 consists almost entirely of transcripts with established 5' terminal oligopyrimidine (TOP) motifs, or, like *Hsp90ab1* and *Ybx1*, with previously unrecognized TOP or related TOP-like motifs that we identified. We find no evidence to support proposals that mTORC1 preferentially regulates mRNAs with increased 5' untranslated region length or complexity<sup>3</sup>. mTORC1 phosphorylates a myriad of translational regulators, but how it controls TOP mRNA translation is unknown<sup>4</sup>. Remarkably, loss of just the 4E-BP family of translational repressors, arguably the best characterized mTORC1 substrates, is sufficient to render TOP and TOP-like mRNA translation resistant to Torin 1. The 4E-BPs inhibit translation initiation by interfering with the interaction between the cap-binding protein eIF4E and eIF4G1. Loss of this interaction diminishes the capacity of eIF4E to bind TOP and TOP-like mRNAs much more than other mRNAs, explaining why mTOR inhibition selectively suppresses their translation. Our results clarify the translational program controlled by mTORC1 and identify 4E-BPs and eIF4G1 as its master effectors.

The mTOR kinase is the catalytic subunit of two complexes, mTORC1 and 2, which regulate growth and are often deregulated in disease (reviewed in ref. 1). mTORC1 is the allosteric target of the well-known drug rapamycin, which has clinical uses in organ transplantation, cardiology and oncology. A major function of mTORC1 is to regulate protein synthesis, which it is thought to control through several substrates, including the S6 kinases (S6Ks), the inhibitory eIF4E-binding proteins (4E-BPs), and the eIF4G initiation factors. ATP-competitive inhibitors of mTOR such as Torin 1 impair protein synthesis and proliferation to a much greater degree than rapamycin<sup>1,2</sup>, largely owing to their inhibition of rapamycin-resistant functions of mTORC1. Because earlier efforts to identify mRNAs translationally regulated by mTORC1 relied on rapamycin<sup>5–7</sup>, it is likely that the mTORC1-regulated translational program is not fully defined.

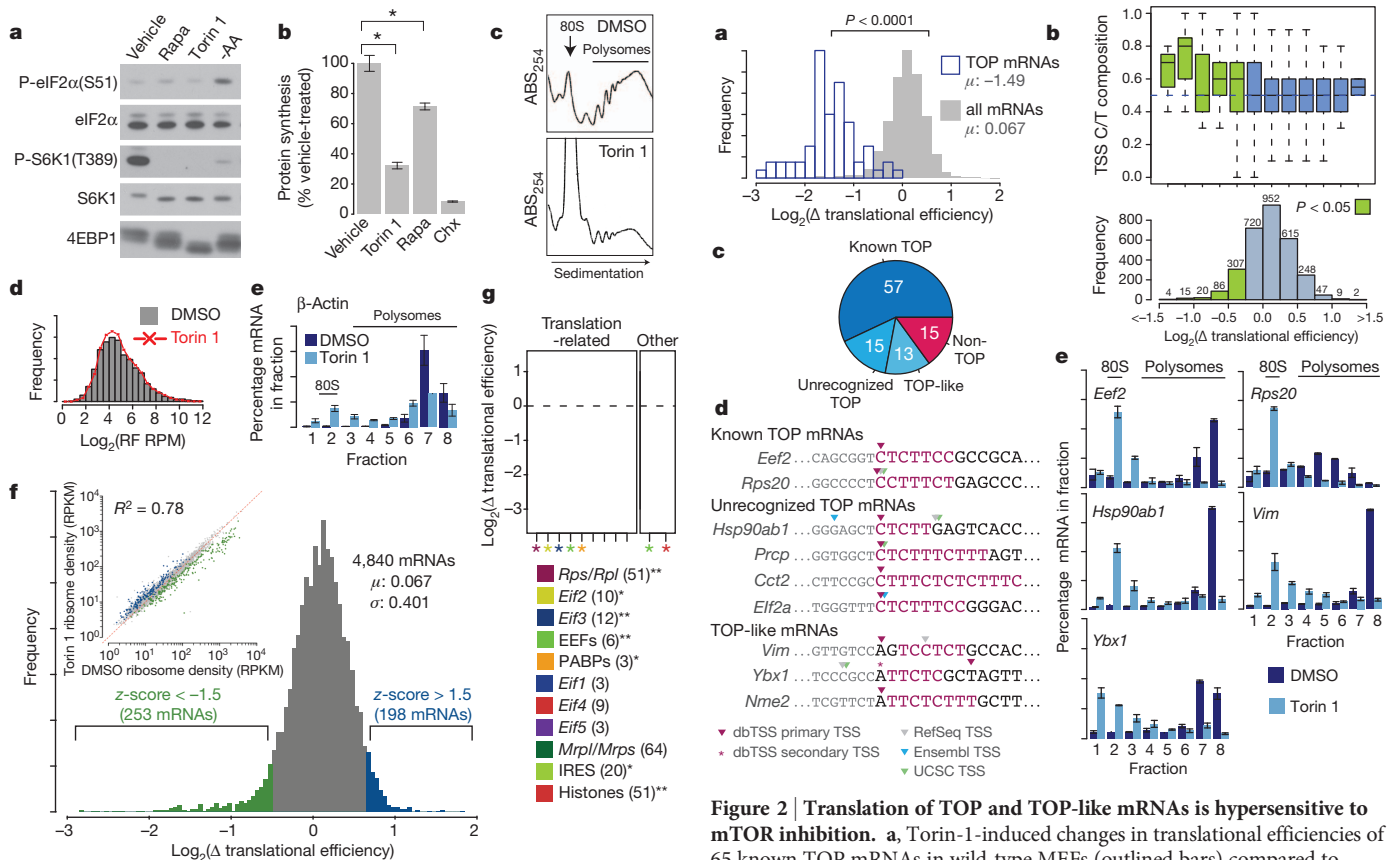
As a step towards defining this program, we examined the effects of Torin 1 on protein synthesis in mouse embryonic fibroblasts (MEFs). To focus on the direct translational outputs of mTORC1 and avoid secondary effects, we treated cells with Torin 1 for only 2 h. Torin 1 blocked canonical mTORC1-dependent events, such as the phosphorylation of S6K1 and

4E-BP1, but did not increase the phosphorylation of eIF2 $\alpha$ , which represses translation and is induced by stresses like amino acid deprivation (Fig. 1a). In wild-type MEFs, Torin 1 suppressed <sup>35</sup>S-Cys/Met incorporation into protein by ~65% and shifted ribosomes out of polysomes, indicating that mTOR inhibition causes a severe defect in translation initiation (Fig. 1b, c).

To monitor systematically the translation of individual mRNAs, we analysed vehicle- and Torin-1-treated MEFs using transcriptome-scale ribosome profiling<sup>8</sup>. Ribosome profiling provides a precise measurement of mRNA translation by quantifying ribosome-protected mRNA fragments (ribosome footprints (RFs)) using deep sequencing. In proliferating MEFs, we detected 3.9 million exon-mapped ribosome footprints that corresponded to 12,856 actively translated RefSeq (<http://www.ncbi.nlm.nih.gov/RefSeq/>) mRNAs. Of these, 4,840 could be monitored at levels sufficient for robust measurements of Torin-1-induced translational changes (Supplementary Table 1). The frequency of RFs that map to each mRNA (gene-specific reads per million total exon-mapped reads, or RPM) reflects the proportion of ribosomes engaged in the translation of that transcript. In vehicle- and Torin-1-treated cells, the distributions of RF frequencies were largely superimposable (median log<sub>2</sub>(change in RF frequency) = 0.08), arguing that mTOR inhibition has similar effects on the translation of most mRNAs (Fig. 1d). Given this and the <sup>35</sup>S-Cys/Met incorporation results (Fig. 1b), we determined (see Methods) that mTOR inhibition suppresses the translation of nearly all (99.8%) mRNAs to some degree, with a mean reduction in translation of 61% (median = 60.5%). Consistent with this conclusion,  $\beta$ -actin mRNA, which, like most mRNAs, underwent little change in RF frequency upon Torin 1 treatment (log<sub>2</sub>( $\Delta$ RPM) = -0.08), was nevertheless partially but significantly depleted from polysomes in Torin-1-treated cells (Fig. 1e). Thus, acute mTOR inhibition has the unappreciated capacity to moderately suppress the translation of nearly all mRNAs.

To identify the mRNAs most regulated by mTOR at the translational level, we calculated the Torin-1-induced change in the translational efficiency of each mRNA (Fig. 1f). This measurement normalizes RF frequency to the abundance of the corresponding transcript and so decouples translational and transcriptional regulation. Using a z-score cut-off of  $\pm 1.5$ , we selected 253 suppressed and 198 resistant mRNAs for further analysis. Gene ontology analyses of Torin-1-suppressed mRNAs showed enrichment for those involved in various steps in protein synthesis (Supplementary Fig. 1a), albeit with differences among components of the translational machinery (Fig. 1g and Supplementary Table 2). For instance, Torin 1 suppressed the translation of *Eif4b* but not of other eIF4F complex components, and of nearly all cytoplasmic ribosomal proteins, except *Rps27a*, which has extra-ribosomal functions<sup>9</sup>. Torin-1-resistant mRNAs are enriched for transcription factors (Supplementary Fig. 1a), such as *Stra13*, *Myc*, *Paf1* and *Foxo1*. Additionally, the translation of mRNAs with putative internal ribosomal entry sites (IRES)<sup>10,11</sup> and, unexpectedly, those encoding histones

<sup>1</sup>Department of Cancer Biology, Dana Farber Cancer Institute, 250 Longwood Avenue, Boston, Massachusetts 02115, USA. <sup>2</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, Massachusetts 02115, USA. <sup>3</sup>Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>4</sup>Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. <sup>5</sup>Broad Institute of Harvard and MIT, Seven Cambridge Center, Cambridge, Massachusetts 02142, USA.



**Figure 1 | Profile of mTOR-regulated translation.** **a**, Wild-type MEFs were treated with vehicle (DMSO), 250 nM rapamycin (Rapa) or Torin 1, or starved for amino acids (–AA) for 2 h and analysed for protein levels. P, phospho. **b**, Wild-type MEFs were treated for 2 h with vehicle (DMSO), 250 nM rapamycin or Torin 1, or 10  $\mu\text{g ml}^{-1}$  cycloheximide (Chx), pulsed for 30 min with  $^{35}\text{S}$ -Cys/Met and  $^{35}\text{S}$  incorporation into protein quantified and normalized to the total protein. Data are mean  $\pm$  s.d. ( $n = 3$ ).  $*P < 0.005$ . **c**, Polysome profiles of wild-type MEFs treated with DMSO or 250 nM Torin 1 for 2 h.  $\text{ABS}_{254}$ , absorbance of light at 254 nm. **d**, Distributions of RF frequency in vehicle- or Torin-1-treated cells. RF libraries from cells treated as in **c** were used to determine RF frequencies (RPM) for 4,840 mRNAs. **e**,  $\beta$ -Actin mRNA abundance in fractions from **c** were quantified by quantitative polymerase chain reaction (qPCR), and calculated as a percentage of the total in all fractions. Data are means  $\pm$  s.e.m. ( $n = 2$ ). **f**, Distribution of changes in translational efficiency from vehicle- or Torin-1-treated cells. RF frequencies from **d** were normalized to transcript levels to calculate translational efficiencies. Ribosome densities (reads per kilobase per million, RPKM) from vehicle- and Torin-1-treated cells are inset. mRNAs with suppressed ( $z\text{-score} < -1.5$ ) or resistant ( $z\text{-score} > 1.5$ ) translational efficiencies are indicated. **g**, Torin-1-dependent changes in translational efficiency for indicated mRNA classes. For histone mRNAs, results reflect changes in ribosome density only. Significance determined by two-tailed Mann–Whitney U test.  $*P < 0.005$ ,  $**P < 0.0005$ .

were also clearly resistant to Torin 1 (Fig. 1g), indicating that these mRNAs use modes of initiation that do not depend on mTOR activity<sup>12</sup>.

We considered the features that define the mRNAs that are most translationally suppressed upon mTOR inhibition. Two types of mRNAs are thought to be highly mTOR-dependent: (1) those with long and complex 5' untranslated regions (UTRs) that are reported to be regulated through a 4E-BP-dependent mechanism<sup>3</sup>; and (2) mRNAs with 5' TOP motifs that are regulated through an unknown mechanism<sup>13</sup>. Surprisingly, the translational efficiency of commonly cited examples of mRNAs with long, complex UTRs, such as cyclin D1 ( $\log_2(\Delta) = -0.07$ ), cyclin D3 ( $\log_2(\Delta) = 0.09$ ), *Myc* ( $\log_2(\Delta) = 0.92$ ) and *Vegfa* ( $\log_2(\Delta) = 0.79$ )<sup>14</sup>, was not significantly suppressed in our data set. We found no evidence that 5' UTR length or complexity correlated positively with sensitivity to mTOR inhibition and, if

**Figure 2 | Translation of TOP and TOP-like mRNAs is hypersensitive to mTOR inhibition.** **a**, Torin-1-induced changes in translational efficiencies of 65 known TOP mRNAs in wild-type MEFs (outlined bars) compared to changes in all 4,840 mRNAs (solid bars). Significance determined by the Mann–Whitney U test. **b**, The pyrimidine content of the 10 nucleotides surrounding the TSS for 3,025 mRNAs where the TSS could be confidently identified, excluding 65 known TOP mRNAs (expected frequency = 0.518). Box plots indicate the TSS pyrimidine content for mRNAs binned according to Torin-1-dependent change in translational efficiency. Significance determined by binomial test. **c**, Numbers of indicated mRNA classes. **d**, TSS annotations for selected TOP and TOP-like mRNAs. Primary and secondary TSS locations from dbTSS (purple) are indicated, as are annotations from RefSeq (grey), Ensembl (blue), and UCSC (green). **e**, Polysome analyses of selected TOP (*Eef2*, *Rps20*), unrecognized TOP (*Hsp90ab1*) and TOP-like (*Vim*, *Ybx1*) mRNAs. Data are means  $\pm$  s.e.m. ( $n = 2$ ).

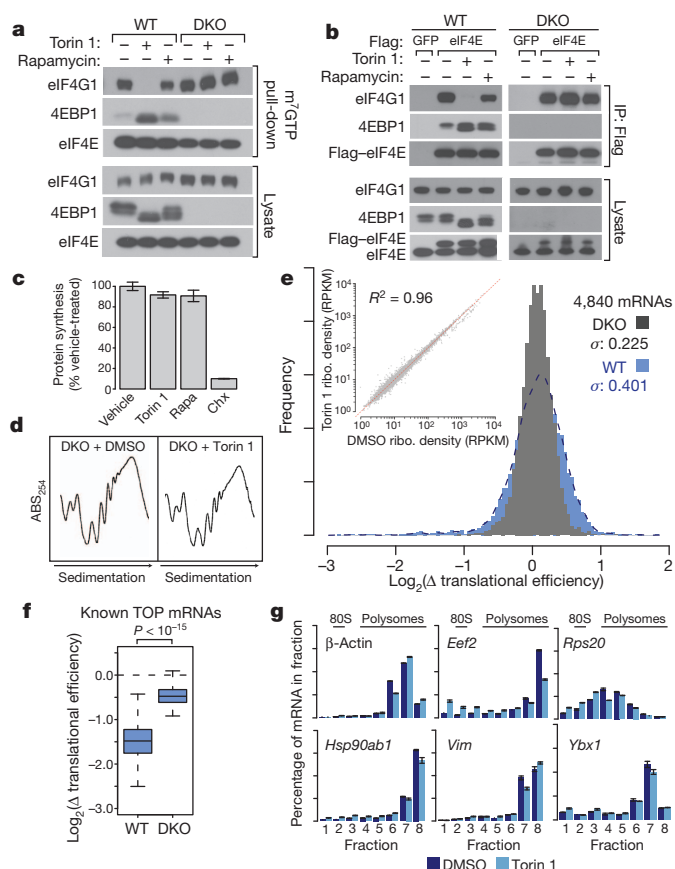
anything, mRNAs with shorter and less complex 5' and 3' UTRs tended to be more sensitive (Supplementary Fig. 1b–d). However, UTR length per se does not determine mTOR dependency because mRNAs with similarly short coding sequence (CDS) and UTR lengths, like those for cytoplasmic and mitochondrial ribosomal proteins (Supplementary Fig. 1b), were differentially sensitive to mTOR inhibition (Fig. 1g). Although it is puzzling that we find little evidence for the selective regulation of mRNAs with complex 5' UTRs, these mRNAs may be affected, upon prolonged mTOR inhibition, by secondary consequences of the acute changes described here. Consistent with this possibility, 24–48 h of mTOR inhibition are required to exclude the cyclin D1 mRNA maximally from polysomes<sup>15,16</sup>.

Torin 1 suppressed the translational efficiencies of all known TOP mRNAs in our data set (mean  $\log_2(\Delta) = -1.49$ ) (Fig. 2a and Supplementary Table 2). TOP mRNAs are defined as those with a cytidine immediately after the 5' cap, followed by an uninterrupted stretch of 4–14 pyrimidines<sup>13,17</sup>, and tend to encode proteins associated with translation<sup>13,18</sup>. When averaged across known TOP mRNAs, Torin 1 depleted RF density throughout the CDS (Supplementary Fig. 2a) and shifted known TOP mRNAs (*Eef2*, *Rps20*) out of polysomes (Fig. 2e). RNA interference (RNAi)-mediated depletion of raptor, an essential mTORC1 component, also selectively inhibited the translation of TOP mRNAs (Supplementary Fig. 3).

Torin 1 also suppressed the translation of many mRNAs not previously defined as TOP mRNAs. After excluding known TOP mRNAs from analysis, we found that the ten nucleotides surrounding the predominant transcriptional start site (TSS) in the mRNAs most suppressed by mTOR inhibition were still highly enriched for pyrimidines (Fig. 2b). This enrichment could reflect the presence of previously undocumented TOP motifs and/or of similar motifs that do not meet the TOP definition. We used the database of transcriptional start sites (dbTSS)<sup>19</sup> as well as the RefSeq, Ensembl and University of California, Santa Cruz (UCSC) resources to examine the TSSs of the 100 mRNAs most translationally suppressed by mTOR inhibition. Fifty-seven of these were known TOP mRNAs, and, of the remaining 43, 15 had previously unrecognized TOP motifs and 13 contained a stretch of pyrimidines that was near but did not begin at the most frequent TSS. As this suggested that the established TOP motif definition might be too conservative, we defined a relaxed TOP-like motif consisting of a stretch of at least five pyrimidines within four nucleotides of the most frequent TSS. Although this motif was relatively common among all TSSs (frequency = 0.16), it was highly enriched among the most suppressed mRNAs and significantly depleted among mRNAs with a greater than average increase in translational efficiency following mTOR inhibition (Fisher's exact test  $P$  value =  $3.1 \times 10^{-8}$ ; Supplementary Fig. 2b). Remarkably, we found that 85 of the 100 mRNAs most sensitive to mTOR inhibition are either known TOP mRNAs or contain an unrecognized TOP or TOP-like motif (Fig. 2c, d and Supplementary Table 3). Several mRNAs that failed to meet our criteria contain pyrimidine sequences interrupted by a single purine (for example, *Hspa8*), suggesting that even our TOP-like definition may be too conservative.

Like established TOP mRNAs, many previously unrecognized TOP and TOP-like mRNAs encode proteins with roles in protein synthesis (Supplementary Table 3) whereas others point to new effectors of the mTORC1 pathway (Fig. 2d). For instance, *Vim* and *Ybx1* participate in the epithelial–mesenchymal transition, a process known to be affected by mTOR inhibition<sup>20,21</sup>. By analysing polysome profiles prepared from Torin-1-treated cells, we confirmed that several unrecognized TOP (*Hsp90ab1*) or TOP-like mRNAs (*Vim*, *Ybx1*) were depleted from polysome fractions as strongly as established TOP mRNAs (*Rps20*, *Eef2*) (Fig. 2e). Furthermore, TOP-like and TOP motifs conferred similar degrees of mTOR-dependent translation control when placed upstream of a luciferase reporter (Supplementary Fig. 4a, b, d). Because some TOP-like mRNAs may be misannotated and actually contain canonical TOP motifs, we *in vitro* transcribed capped mRNA beginning with a single purine followed by a pyrimidine sequence and found that, like TOP mRNAs, it was translated less efficiently than an mRNA lacking this motif when mTOR was inhibited (Supplementary Fig. 4e, f). Thus, TOP and TOP-like motifs are more numerous than previously recognized and define the vast majority of mRNAs highly dependent on mTOR for translation.

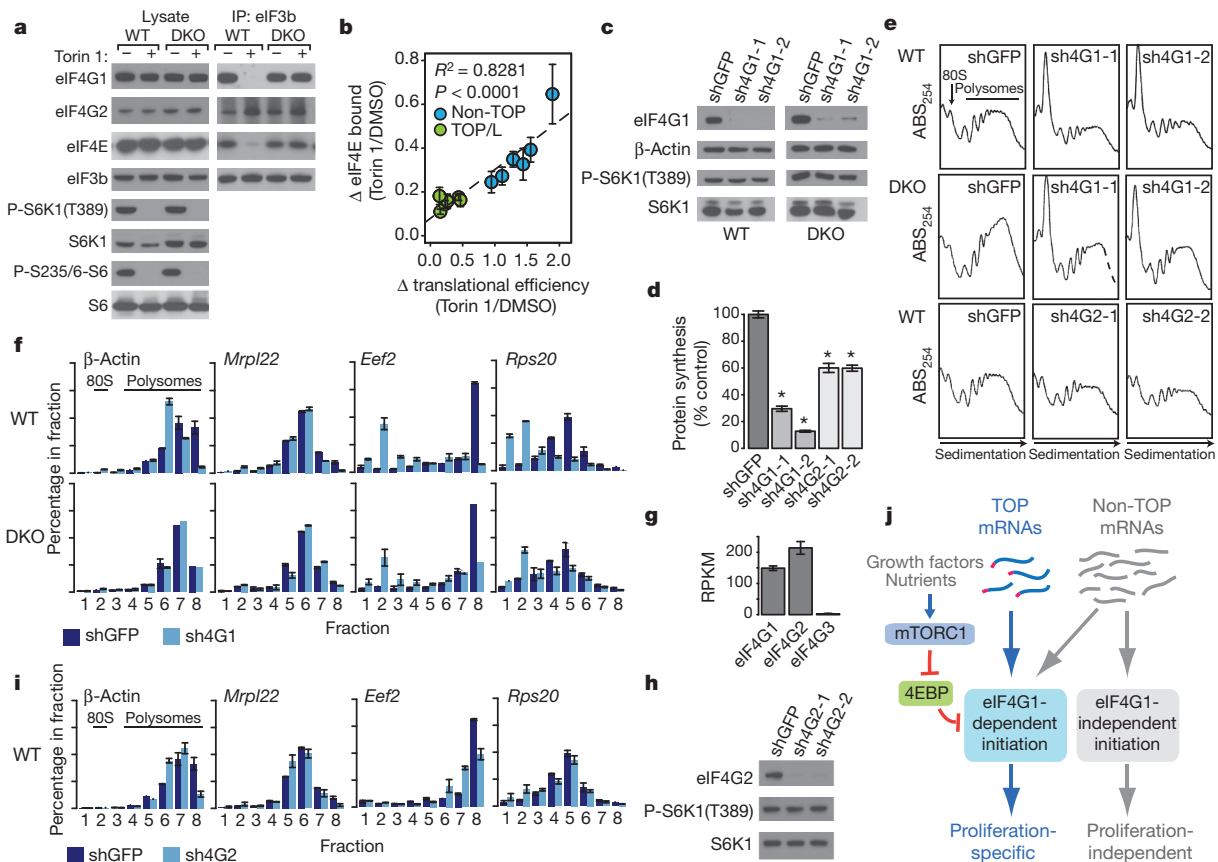
How mTOR regulates TOP mRNA translation has been a persistent mystery. The S6Ks were originally considered key mediators, but later studies did not support this possibility<sup>22,23</sup>. Because TOP mRNA translation is less inhibited by rapamycin than dual mTOR/PI3K inhibitors and RNAi-mediated mTOR suppression<sup>4</sup>, we suspected that it might be regulated through the 4E-BPs, which mTORC1 phosphorylates in a largely rapamycin-resistant fashion<sup>2,24,25</sup>. In 4E-BP1 and 4E-BP2 (also known as Eif4ebp1 and 2, respectively) double-knockout MEFs (DKO), Torin 1 had no effect on the interaction of eIF4E with eIF4G1 (Fig. 3a, b). Furthermore, in DKO cells, Torin 1 had a minimal effect on <sup>35</sup>S-Cys/Met incorporation and did not perceptibly shift ribosomes out of polysomes (Fig. 3c, d), indicating that the 4E-BPs mediate a large part of mTOR-dependent control of general translation. Moreover, ribosome profiling of vehicle- and Torin-1-treated DKO cells revealed that the distribution of Torin-1-induced changes in translational efficiency was much narrower in DKO ( $\sigma = 0.225$ ) than in wild-type ( $\sigma = 0.401$ ) cells (Fig. 3e), indicating that the 4E-BPs are



**Figure 3 | mTOR regulates general protein synthesis and TOP mRNA translation through the 4E-BPs.** **a**, Wild-type (WT) and 4EBP1/2 DKO MEFs were treated with DMSO, 250 nM rapamycin or Torin 1 for 2 h, lysates were subjected to m<sup>7</sup>GTP pull-downs, and analysed for levels of indicated proteins. **b**, Wild-type and DKO MEFs expressing Flag–GFP or Flag–eIF4E were treated as in **a**, and immunoprecipitates (IP) analysed for indicated proteins. **c**, DKO MEFs were treated for 2 h with vehicle (DMSO), 250 nM rapamycin (Rapa) or Torin 1, or 10 µg ml<sup>−1</sup> cycloheximide (Chx) were analysed as in Fig. 1b. Data are mean ± s.d. ( $n = 3$ ). **d**, Polysome profiles of DKO MEFs treated with DMSO or Torin 1 for 2 h. **e**, Torin-1-dependent changes in translational efficiency in DKO (grey bars) and wild-type MEFs (blue bars). **f**, Torin-1-dependent translational suppression of 65 TOP mRNAs in wild-type and DKO MEFs. Significance determined by Mann–Whitney U test. **g**, Polysome analyses of selected non-TOP ( $\beta$ -actin), known TOP (*Eef2*, *Rps20*), unrecognized TOP (*Hsp90ab1*) and TOP-like (*Vim*, *Ybx1*) mRNAs in DKO cells. Data are means ± s.e.m. ( $n = 2$ ).

also required for the largest translational effects caused by mTOR inhibition. Indeed, as monitored by ribosome profiling, established TOP mRNAs were barely inhibited by Torin 1 in DKO cells (Fig. 3f), which we confirmed by polysome analysis of individual mRNAs in MEFs (Fig. 3g) and in HeLa cells with RNAi-mediated knockdown of 4E-BP1 (Supplementary Fig. 5). Expression of a dominant-negative 4EBP1-4A mutant, as well as RNAi-mediated depletion of eIF4E, were sufficient to inhibit TOP mRNA translation selectively in actively growing cells (Supplementary Fig. 6). Expression of the 4EBP1-4A mutant suppressed the translation of TOP reporter constructs as well (Supplementary Fig. 4c). We found no evidence that previously identified pyrimidine-binding proteins, such as TIA1, TIAR or La (also known as SSB), have a role in the selective regulation of TOP mRNAs by mTORC1 (Supplementary Fig. 7). However, we cannot rule out a role for these proteins in the amino acid regulation of TOP mRNA translation, which is maintained in DKO cells probably through the GCN2 (also known as eIF2AK4) pathway (Supplementary Fig. 8). These results indicate that the translation of mRNAs with TOP and TOP-like motifs is highly sensitive to 4E-BP phosphorylation, and that this is the basis of their regulation by mTORC1.





**Figure 4 | Destabilization of the eIF4E–eIF4G1 interaction dissociates TOP mRNAs from eIF4E and inhibits their translation.** **a**, Wild-type (WT) and DKO MEFs were treated for 2 h with DMSO or 250 nM Torin 1 and eIF3b immunoprecipitates analysed for indicated proteins. P, phospho. P-S235/236-S6 indicates S6 phosphorylated at Ser 235 and Ser 236. **b**, Flag-eIF4E was immunoprecipitated from wild-type MEFs treated with DMSO or 250 nM Torin 1 for 2 h. RNA was extracted, and the abundance of TOP and TOP-like (TOP/L) (*Eef2*, *Rps20*, *Hsp90ab1*, *Pabpc1*, *Ybx1*, *Vim*) and non-TOP (*Actb*, *Mrpl22*, *Ccnd1*, *Slc2a1*, *Gabarp1*, *Myc*) mRNAs was quantified by qPCR. Changes in eIF4E binding of mRNAs were plotted against changes in translational efficiency from Fig. 1f. eIF4E binding data are means  $\pm$  s.e.m. ( $n = 4$ ). **c**, Levels of indicated proteins in cells expressing indicated short hairpin RNAs (shRNAs). **d**, Cells

To understand why the translation of TOP and TOP-like mRNAs has a 4E-BP-mediated hyper-dependence on mTOR, we considered the established functions of the 4E-BPs<sup>3</sup>. A key step in eIF4E-dependent initiation is the cooperative binding of eIF4E and eIF4G1 to mRNA, which nucleates the eIF4F complex<sup>26</sup>. eIF4G1 also interacts with eIF3, which orchestrates assembly of the 43S pre-initiation complex on the mRNA. When mTORC1 is inactive, dephosphorylated 4E-BPs bind to eIF4E and thereby prevent its association with eIF4G1 (Fig. 3a, b). mTOR inhibition also prevents the association of eIF4G1 with eIF3 in wild-type but, unexpectedly, not in DKO cells<sup>27</sup> (Fig. 4a). Expression of the 4EBP1-4A mutant similarly disrupted the eIF4G1–eIF3 interaction (Supplementary Fig. 6b). Because destabilization of the eIF4F complex weakens the affinity of eIF4E for the mRNA cap<sup>26</sup>, we hypothesized that mTOR inhibition might selectively impair the binding of eIF4E to TOP and TOP-like mRNAs. Indeed, Torin 1 treatment of cells caused a selective loss of TOP and TOP-like mRNAs from eIF4E, which strongly correlated with their degree of translational suppression (Fig. 4b). Consistent with a special role for eIF4G1 in TOP mRNA translation, RNAi-mediated depletion of eIF4G1 in wild-type cells, which mimicked the effects of Torin 1 on overall protein synthesis and polysome profiles, selectively suppressed the translation of TOP mRNAs, without affecting mTORC1 activity (Fig. 4c–f). Importantly, in the DKO cells, eIF4G1 depletion also selectively repressed TOP

expressing indicated shRNAs were pulsed for 30 min with <sup>35</sup>S-labelled Cys/Met and analysed as in Fig. 1b. Data are means  $\pm$  s.d. ( $n = 3$ ). Significance was determined by *t*-test. \* $P < 0.001$ . **e**, Polysome profiles for wild-type or DKO cells expressing indicated shRNAs. **f**, RNA isolated from gradients in **e** was analysed by qPCR for the indicated mRNAs as in Fig. 1e. Data are means  $\pm$  s.e.m. ( $n = 2$ ). **g**, Abundance of indicated transcripts from RNA-Seq analysis. Data are means  $\pm$  s.e.m. ( $n = 3$ ). **h**, Lysates from cells expressing shGFP or eIF4G2-specific shRNAs were analysed by immunoblotting. **i**, Fractions from shEIF4G2-2 gradients in **e** were analysed as in **f**. **j**, mTORC1 regulates the selective translation of TOP and TOP-like mRNAs through the 4EBP-dependent control of eIF4G1-mediated initiation.

mRNA translation (Fig. 4c, e, f), consistent with eIF4G1 acting downstream of the 4E-BPs. A functionally redundant eIF4G1 homologue, eIF4G3, is not well expressed in the MEFs (Fig. 4g) and its loss had little effect on translation in HeLa cells (Supplementary Fig. 9). MEFs do express a distinct eIF4G1 homologue, eIF4G2 (also known as DAP5; Fig. 4g), which does not bind eIF4E but still mediates a substantial fraction of protein synthesis<sup>28,29</sup>. Although eIF4G2 depletion significantly suppressed overall protein synthesis, it did not have selective effects on the translation of TOP mRNAs (Fig. 4a, d, e, h, i). Therefore, unlike other mRNAs, TOP mRNAs require eIF4G1 to anchor eIF4E to the cap, and this underlies their selective translational regulation by the 4E-BPs and mTORC1.

We find that the effects of acute mTOR inhibition on mRNA translation are largely mediated by the 4E-BPs, including the moderate suppression of the translation of all mRNAs and the more marked inhibition of TOP and TOP-like mRNA translation. As the 4E-BPs are required for the mTORC1-dependent regulation of proliferation<sup>16</sup>, the translational control of TOP mRNAs may have a fundamental role in this process (Fig. 4j), as well as in cancers associated with hyperactive mTOR signalling. We focused on suppressed mRNAs, but many other transcripts are translated with increased efficiency, and may be important for cellular survival under conditions of impaired mTORC1 signalling.

## METHODS SUMMARY

To generate ribosome and mRNA profiling libraries, wild-type MEFs (4EBP1/2<sup>+/+</sup>; p53<sup>-/-</sup>) or DKO MEFs (4EBP1/2<sup>-/-</sup>; p53<sup>-/-</sup>) were treated with vehicle or 250 nM Torin 1 for 2 h. Cellular extracts were partitioned for either ribosome profiling or mRNA profiling. Small RNA libraries were prepared according to established protocols<sup>8</sup> with some modifications, and analysed by high-throughput sequencing. Transcript abundance was determined through an iterative alignment and mapping strategy to a non-redundant library of mouse transcripts based on RefSeq definitions.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 30 June 2011; accepted 29 March 2012.**

- Zoncu, R., Efeyan, A. & Sabatini, D. M. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nature Rev. Mol. Cell Biol.* **12**, 21–35 (2011).
- Thoreen, C. C. *et al.* An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J. Biol. Chem.* **284**, 8023–8032 (2009).
- Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev.* **18**, 1926–1945 (2004).
- Patursky-Polischuk, I. *et al.* The TSC-mTOR pathway mediates translational activation of TOP mRNAs by insulin largely in a raptor- or rictor-independent manner. *Mol. Cell. Biol.* **29**, 640–649 (2009).
- Bilanges, B. *et al.* Tuberous sclerosis complex proteins 1 and 2 control serum-dependent translation in a TOP-dependent and -independent manner. *Mol. Cell. Biol.* **27**, 5746–5764 (2007).
- Gera, J. F. *et al.* AKT activity determines sensitivity to mammalian target of rapamycin (mTOR) inhibitors by regulating cyclin D1 and *c-myc* expression. *J. Biol. Chem.* **279**, 2737–2746 (2004).
- Grolleau, A. *et al.* Global and specific translational control by rapamycin in T cells uncovered by microarrays and proteomics. *J. Biol. Chem.* **277**, 22175–22184 (2002).
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
- Redman, K. L. & Rechsteiner, M. Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* **338**, 438–440 (1989).
- Mokrejs, M. *et al.* IRESite—a tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Res.* **38**, D131–D136 (2010).
- Kozak, M. A second look at cellular mRNA sequences said to function as internal ribosome entry sites. *Nucleic Acids Res.* **33**, 6593–6602 (2005).
- Martin, F. *et al.* Cap-assisted internal initiation of translation of histone H4. *Mol. Cell* **41**, 197–209 (2011).
- Meyuhas, O. Synthesis of the translational apparatus is regulated at the translational level. *Eur. J. Biochem.* **267**, 6321–6330 (2000).
- De Benedetti, A. & Graff, J. R. eIF-4E expression and its role in malignancies and metastases. *Oncogene* **23**, 3189–3199 (2004).
- Shi, Y., Sharma, A., Wu, H., Lichtenstein, A. & Gera, J. Cyclin D1 and *c-myc* internal ribosome entry site (IRES)-dependent translation is regulated by AKT activity and enhanced by rapamycin through a p38 MAPK- and ERK-dependent pathway. *J. Biol. Chem.* **280**, 10964–10973 (2005).
- Dowling, R. J. *et al.* mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328**, 1172–1176 (2010).
- Jefferies, H. B., Reinhard, C., Kozma, S. C. & Thomas, G. Rapamycin selectively represses translation of the “polypyrimidine tract” mRNA family. *Proc. Natl Acad. Sci. USA* **91**, 4441–4445 (1994).
- Iadevaia, V., Caldarola, S., Tino, E., Amaldi, F. & Loreni, F. All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. *RNA* **14**, 1730–1736 (2008).
- Yamashita, R., Wakaguri, H., Sugano, S., Suzuki, Y. & Nakai, K. DBTSS provides a tissue specific dynamic view of transcription start sites. *Nucleic Acids Res.* **38**, D98–D104 (2010).
- Lamouille, S. & Derynck, R. Cell size and invasion in TGF- $\beta$ -induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *J. Cell Biol.* **178**, 437–451 (2007).
- Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
- Jefferies, H. B. *et al.* Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k. *EMBO J.* **16**, 3693–3704 (1997).
- Pende, M. *et al.* S6K1<sup>-/-</sup>/S6K2<sup>-/-</sup> mice exhibit perinatal lethality and rapamycin-sensitive 5'-terminal oligopyrimidine mRNA translation and reveal a mitogen-activated protein kinase-dependent S6 kinase pathway. *Mol. Cell. Biol.* **24**, 3112–3124 (2004).
- Choo, A. Y., Yoon, S. O., Kim, S. G., Roux, P. P. & Blenis, J. Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc. Natl Acad. Sci. USA* **105**, 17414–17419 (2008).
- Feldman, M. E. *et al.* Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biol.* **7**, e38 (2009).
- Ptushkina, M. *et al.* Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *EMBO J.* **17**, 4798–4808 (1998).
- Harris, T. E. *et al.* mTOR-dependent stimulation of the association of eIF4G and eIF3 by insulin. *EMBO J.* **25**, 1659–1668 (2006).
- Lee, S. H. & McCormick, F. p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. *EMBO J.* **25**, 4008–4019 (2006).
- Ramirez-Valle, F., Braunstein, S., Zavadii, J., Formenti, S. C. & Schneider, R. J. eIF4G links nutrient sensing by mTOR to cell proliferation and inhibition of autophagy. *J. Cell Biol.* **181**, 293–307 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank members of the Gray and Sabatini laboratories for helpful discussions, H. Guo, S. Hawthorne, G. Brar, J. Damon, C. Miller and W. Gilbert for advice and N. Sonenberg for providing 4EBP1/2 wild-type and double-knockout MEFs. This work was supported by the National Institutes of Health (CA103866 and CA129105 to D.M.S.), Department of Defense (W81XWH-07-0448 to D.M.S.), the W.M. Keck Foundation (D.M.S.), LAM Foundation (D.M.S.), Dana Farber Cancer Institute (N.S.G., C.C.T.), and fellowship support from the American Cancer Society (C.C.T.), and the National Science Foundation (L.C. and T.W.). D.M.S. is an investigator of the Howard Hughes Medical Institute.

**Author Contributions** C.C.T. and D.M.S. conceived the project. C.C.T. designed and performed most experiments and data analyses with input from D.M.S. and N.S.G. L.C. and H.R.K. assisted with experiments and T.W. with sequence analysis. C.C.T. and D.M.S. wrote and edited the manuscript with input from N.S.G.

**Author Information** Small RNA sequencing data were deposited in the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession number GSE36892. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to N.S.G. (nathanael\_gray@dfci.harvard.edu) or D.M.S. (sabatini@wi.mit.edu).

## METHODS

**Materials.** Reagents were obtained from the following sources: antibodies to phospho-Thr-389 S6K, S6, 4EBP1, eIF4E, phospho-S51 eIF2 $\alpha$ , eIF2 $\alpha$ , eIF4G1 and eIF4G2 from Cell Signaling; antibodies to eIF3b (N20),  $\beta$ -actin, S6K and horseradish-peroxidase-labelled anti-mouse, anti-goat and anti-rabbit secondary antibodies from Santa Cruz Biotechnology; FuGENE 6 and Complete Protease Mixture from Roche Applied Science; Flag M2-agarose and cycloheximide from Sigma; 7-methyl-GTP-Sepharose from GE Healthcare; rapamycin from LC Laboratories; luciferase mRNA and luciferase assay reagents from Promega; EasyTagTM EXPRESS  $^{35}$ S protein labelling mix from PerkinElmer Life Sciences; acid phenol/chloroform/isoamyl alcohol, GlycoBlue, Superscript In and Proteinase K from Ambion. Polynucleotide kinase (PNK), polyA polymerase, Phusion High Fidelity DNA polymerase, and RNase I, T7 RNA polymerase, DNase I and m7GpppG cap analogue from New England Biolabs; CircLigase from Epicentre Biotechnologies; Ecl136 from Fermentas; DMEM from SAFC Biosciences; TransMessenger transfection reagent from Qiagen; inactivated fetal calf serum and oligo dT<sub>25</sub> Dynabeads from Invitrogen. 4EBP1/2 $^{+/+}$ ; p53 $^{-/-}$  MEFs and 4EBP1/2 $^{-/-}$ ; p53 $^{-/-}$  MEFs were provided by N. Sonenberg (McGill University). Torin 1 was synthesized and purified in the Gray laboratory<sup>30</sup> and is commercially available from Tocris.

**Preparation of cell lysates and affinity purifications.** Cells were rinsed once with ice-cold PBS and lysed in ice-cold lysis buffer (buffer A: 50 mM HEPES-KOH (pH 7.4), 2 mM EDTA, 10 mM pyrophosphate, 10 mM  $\beta$ -glycerophosphate, 40 mM NaCl, 1% Triton X-100 and one tablet of EDTA-free protease inhibitors (Roche) per 25 ml). The soluble fractions of cell lysates were isolated by centrifugation at 13,000g for 10 min. For immunoprecipitations, primary antibodies were added to lysates and incubated with rotation for 2 h at 4 °C. 20  $\mu$ l of a 50% slurry of protein G-sepharose was then added and the incubation continued for an additional 1 h. Immunoprecipitates were washed three times with lysis buffer. Immunoprecipitated proteins were denatured by the addition of 20  $\mu$ l of sample buffer and boiling for 5 min, resolved by 8–16% SDS-PAGE, and analysed by western blot. For Flag purifications, Flag M2 affinity gel was washed with lysis buffer three times. 20  $\mu$ l of beads in 50% slurry was then added to cell lysates and incubated with rotation for 2 h at 4 °C. Finally, beads were washed three times with lysis buffer. Immunoprecipitated proteins were denatured by the addition of 50  $\mu$ l of sample buffer and boiling for 5 min. For m7GTP affinity purifications, m7GTP sepharose was washed with lysis buffer. 20  $\mu$ l of beads in 50% slurry was added to cell lysates and incubated with rotation for 2 h at 4 °C. Finally, beads were washed three times with lysis buffer, denatured by the addition of 50  $\mu$ l sample buffer and analysed by western blot. For amino acid starvation, cells were washed twice in amino acid-free RPMI and then incubated in RPMI containing 10% dialysed FBS with or without amino acids.

**Metabolic labelling of cells.** Cells were seeded in 6-well plates and cultured overnight. Cells were then treated with appropriate compounds for 2 h, washed once with cysteine/methionine-free DMEM, and then incubated in 2 ml of cysteine/methionine-free DMEM, 10% dialysed inactivated fetal calf serum, compound, and 165  $\mu$ Ci (15  $\mu$ l, 11  $\mu$ Ci  $\mu$ l<sup>-1</sup>) of EasyTag EXPRESS  $^{35}$ S protein labelling mix. After 30 min, cells were lysed, and soluble fractions were isolated by centrifugation at 13,000g for 10 min. Lysates were then spotted on Whatman filter paper and protein was precipitated with 5% trichloroacetic acid, washed two times for 5 min in cold 10% trichloroacetic acid, washed two times for 2 min in cold ethanol, washed one time for 2 min in acetone, and air-dried at room temperature (25 °C). The amount of  $^{35}$ S incorporated into protein was measured using a Beckman LS6500 Scintillation Counter. Total protein content was determined by Bradford assay (Bio-rad).

**Mammalian lentiviral shRNAs and cDNAs.** All shRNA vectors were obtained from the collection of The RNAi Consortium at the Broad Institute<sup>31</sup>. These shRNAs are named with the numbers found at the RNAi Consortium public website: mouse *shEif4g1-1* (TRCN0000096809), mouse *shEif4g1-2* (TRCN0000096811), mouse *shEif4g2-1* (TRCN000009807), mouse *shEif4g2-2* (TRCN000009809), mouse *shEif4e-1* (TRCN0000077475), mouse *shEif4e-2* (TRCN0000077477), mouse *shTia1* (TRCN0000077161), mouse *shTiar* (TRCN0000102619), mouse shRaptor (TRCN0000077472), human *sh4EBP1* (TRCN0000040203), human *sh4EBP2* (TRCN0000117814), human *shEIF4G1-1* (TRCN0000061769), human *shEIF4G1-2* (TRCN0000061770), human *shEIF4G3-1* (TRCN0000142702), human *shEIF4G3-2* (TRCN0000139543). shRNA-encoding plasmids were co-transfected with the psPax2 envelope and vesicular stomatitis virus G packaging plasmids into actively growing HEK-293T using FuGENE 6 transfection reagent as described previously<sup>32</sup>. Virus-containing supernatants were collected at 48 h after transfection and filtered to eliminate cells, and target cells were infected in the presence of 8  $\mu$ g ml<sup>-1</sup> polybrene. Twenty-four hours later, cells were selected with puromycin and analysed on the fourth day after infection. 4EBP1-4A mutant was constructed by mutating T36,

T47, S65 and T70 residues of rat 4EBP1 to alanines, which was then inserted into a Tet-On plasmid.

**Polysome analysis, RNA isolation, and qPCR.** Cells were seeded in 15 cm dishes at  $5 \times 10^6$  cells per dish and cultured overnight. Cells were then treated with 100  $\mu$ g ml<sup>-1</sup> cycloheximide for 5 min before lysis, washed in ice-cold PBS—(PBS lacking calcium) plus 100  $\mu$ g ml<sup>-1</sup> cycloheximide, and then lysed in polysome lysis buffer (15 mM HEPES-KOH (pH 7.4), 7.5 mM MgCl<sub>2</sub>, 100 mM KCl, 2 mM DTT, 1.0% Triton X-100, 100  $\mu$ g ml<sup>-1</sup> cycloheximide, and one tablet of EDTA-free protease inhibitors (Roche) per 25 ml). Lysates were normalized by protein content using Bradford reagent (Bio-rad) and either layered onto 11 ml 10–50% sucrose density gradients (15 mM HEPES-KOH, 7.5 mM MgCl<sub>2</sub>, 100 mM KCl, 2 mM DTT, 100  $\mu$ g ml<sup>-1</sup> cycloheximide, 20 U ml<sup>-1</sup> Superscript In, 10–50% RNase-free sucrose) or adjusted to 0.5% SDS and reserved for total RNA isolation. Gradients were centrifuged in an SW-41Ti rotor at 32,000 r.p.m. at 4 °C for 2 h, and then sampled using a Labconco Auto Densi-Flow Gradient Fractionator connected to an Isco Tris pump with constant monitoring of optical density (OD) at 254 nm. 1 ml fractions were collected throughout, adjusted to 0.5% SDS and incubated at 65 °C for 5 min. 5 ng of polyA+ synthetic luciferase mRNA (Promega) was added to each fraction for normalization. Samples were then treated with 200  $\mu$ g ml<sup>-1</sup> Proteinase K (Ambion) and digested for 45 min at 50 °C, followed by 1:1 dilution with RNase-free water. RNA was extracted from diluted fractions using the hot acid phenol method, and precipitated with NaOAc and isopropanol. cDNA was prepared using the Superscript III reverse transcriptase (Invitrogen) with random hexamer primers according to the manufacturer's instructions. Transcript abundance was determined by quantitative PCR (qPCR) using SYBR Green PCR mix (Applied Biosystems) and primers specific for each transcript. Measurements were then normalized to luciferase abundance, and plotted as per cent detected.

Oligonucleotides used for qPCR of mouse mRNAs are as follows. *Eef2*: forward 5'-GAGAATCCGTCGCCATCCGCCA-3', reverse 5'-CGGGCTTGATGCGTT CAGCGA-3';  $\beta$ -actin: forward 5'-TCGTTGCCGGTCCACACCCG-3', reverse 5'-CTCCTCAGGGGCCACACGCAG-3'; *Mrlp21*: forward 5'-TCTGGGCAAC GCAGACGCTG-3', reverse 5'-GCCAAGCGACCTCGGCCAT-3'; *Rps20*: forward 5'-TGACTCACCCTGTTGCTCC-3', reverse 5'-GAGTCGCTTGTTG ATCCTCATCTGG-3'; *Hsp90ab1*: forward 5'-GCCGTGCGAGTCGGACT TGGT-3', reverse 5'-CCGACACCAAACTGCCCGATCA-3'; *Vim*: forward 5'-ACTGCTGCCCTGCGTGATGT-3', reverse 5'-TCTCAGCATCTGGCGCT CC-3'; *Ybx1*: forward 5'-GGGGTCTCCACGCAATTACC-3', reverse 5'-CG GCGATACCGACGTTGAGGT-3'; *Pabpc1*: 5'-CGCTGGACTGCTCAGGGT GC-3', reverse 5'-GGGGGCGCAGATGCCAACAT-3'; *Myc*: forward 5'-GCC AGCCCTGAGCCCTAGT-3', reverse 5'-GGGTGCGCGGTAGTTGTGCT-3'; *Gabara1*: 5'-AGCCCCCAAAGCTCGGATAGGA-3', reverse 5'-GGTGTTC TGGTACAGCTGACCC-3'; *Slc2a1*: forward 5'-CTGGCATGGCAGGCT GTGCT-3', reverse 5'-CGCCCCCAGAGGGTGAAGA-3'; *Ccnd1*: forward 5'-GCCCGAGGAGCTGCTGCAAA-3', reverse 5'-GCCTTGCATCGCAGCC ACCA-3'; firefly luciferase: forward 5'-ATCCGGAAGCGACCAACGCC-3', reverse 5'-GTCGGGAAGACCTGCCACGC-3'.

Oligonucleotides used for qPCR of human mRNAs are as follows.  $\beta$ -Actin: forward 5'-AGCCTCGCCTTTGCCGA-3', reverse 5'-GCGCGGCGATATCA TCATC-3'; *Gnb2l*: forward 5'-TGGGATCTCACACGGGCACCA-3', reverse 5'-CCGGTTGTGAGAGGAGAAGGCCA-3'; *Rps20*: forward 5'-CCAGTTCGA ATGCATACCAAGACTT-3', reverse 5'-ACTTCCACCTCAACTCTGGCT CA-3'; *eIF4G3*: forward 5'-CCAGAGGGCTGCCTCTATCA-3', reverse 5'-TGGCAATCCATGCTGCTCTGC-3'.

**Protein–RNA co-immunoprecipitation assays.** MEFs stably expressing the indicated Flag-tagged constructs were seeded in 10 cm plates at  $2 \times 10^6$  cells per plate and incubated overnight. Cells were then treated with vehicle or 250 nM Torin 1 for 2 h, and lysed in buffer A (see earlier) containing 40 U ml<sup>-1</sup> Superscript In. Insoluble material was removed by centrifugation and lysates were normalized by protein concentration and incubated with Flag-M2 agarose for 2 h at 4 °C with rotation. Immunoprecipitates were then washed six times with 1 ml buffer A, twice with polysome lysis buffer, and eluted with 100  $\mu$ l 3 $\times$ Flag peptide in polysome lysis buffer for 10 min at 37 °C. Eluates were divided into portions for immunoblotting and RNA extraction. For RNA extractions, 10 ng luciferase mRNA, 1  $\mu$ g yeast tRNA, 200  $\mu$ g ml<sup>-1</sup> proteinase K and SDS (0.5% final concentration) were added to eluates, which were then incubated at 50 °C for 45 min. RNA was extracted twice with acid phenol, once with chloroform and precipitated with NaOAc and isopropanol. Isolated RNA was used as a template for cDNA synthesis using oligo-dT primers, and analysed by qPCR. mRNA abundance in each sample was normalized to spike-in luciferase. For western blotting, sample buffer was added to eluates, which were analysed as described earlier.

**Luciferase reporter assays.** For luciferase reporter assays, plasmids were constructed by cloning the 5' UTRs and 1 kb of upstream sequence into a derivative



of the pIS1 renilla luciferase expression vector where the CMV promoter had been excised. The pIS1 3' UTR was maintained intact. The *Vim/Actb* hybrid 5' UTR reporter was constructed by replacing the promoter and first 30 nucleotides of the *Actb* reporter with the promoter and first 30 nucleotides of the *Vim* reporter. Cells were then seeded in 6-well plates at  $10^5$  cells per well and simultaneously transfected with 100 ng of the indicated reporter plasmid and 400 ng of empty vector. After an overnight incubation, cells were washed with fresh media and treated with either vehicle or 250 nM Torin 1 for 24 h. Luciferase expression was quantified using the Renilla Luciferase Assay System (Promega) according to the manufacturer's directions using a standard laboratory luminometer.

For mRNA transfection experiments, *Eef2* and *Actb* 5' UTRs immediately preceded by a T7 RNA polymerase promoter were cloned into a derivative of pRL containing a renilla luciferase open reading frame (ORF), a short 3' UTR and poly(A)<sub>62</sub> tail followed by an *Ecl*136 restriction site. 50 µg of each reporter plasmid were digested overnight with *Ecl*136, purified by phenol-chloroform extraction, and used as a template for T7 RNA polymerase in reactions containing a fivefold excess of m7GpppG cap analogue. mRNA was then purified by acid phenol extraction. For transfection, cells were seeded in 24-well plates at 50,000 cells per well and transfected with 200 ng *Eef2* or *Actb* renilla luciferase reporter mRNA and 200 ng of a control firefly luciferase mRNA for 2 h using the TransMessenger mRNA transfection reagent according to the manufacturer's directions. Cells were then washed with fresh media containing serum, incubated for 1 h, and then treated with vehicle or 250 nM Torin 1. After a 16 h incubation, luciferase production was quantified using the Dual Luciferase Reporter Assay (Promega) according to the manufacturer's directions. Renilla expression values were then normalized to firefly expression values to control for transfection efficiency.

**Small RNA library preparation.** Footprint libraries were prepared as described previously with minor modifications<sup>8</sup>. Briefly, cells were seeded in 15 cm dishes at  $5 \times 10^6$  cells per plate and cultured overnight. Importantly, we ensured that cells had not reached confluency by the following day, as confluency is known to significantly affect mRNA translation<sup>33</sup>. Cells were then treated with vehicle (DMSO), rapamycin or Torin 1 for 2 h. Five minutes before lysis, 100 µg ml<sup>-1</sup> cycloheximide was added to each plate. Cells were then washed once with ice-cold PBS— plus 100 µg ml<sup>-1</sup> cycloheximide, and lysed in footprint lysis buffer (15 mM HEPES-KOH (pH 7.4), 7.5 mM MgCl<sub>2</sub>, 300 mM KCl, 100 µg ml<sup>-1</sup> cycloheximide, 2 mM DTT, 1.0% Triton X-100, 1 tablet of EDTA-free protease inhibitors (Roche) per 25 ml). Lysates were cleared by centrifugation at 13,000g at 4 °C for 10 min, and supernatants were transferred to clean tubes. RNase I<sub>f</sub> was added to a final concentration of 1 U µl<sup>-1</sup>, and samples were incubated at 25 °C for 45 min with constant rotation. Digested samples were then layered onto 10–50% sucrose density gradients (15 mM HEPES-KOH (pH 7.4), 7.5 mM MgCl<sub>2</sub>, 300 mM KCl, 2 mM DTT, 100 µg ml<sup>-1</sup> cycloheximide, 20 U ml<sup>-1</sup> SupersaseIn, 10–50% RNase-free sucrose) and centrifuged in an SW-41Ti rotor at 36,000 r.p.m. for 2.5 h. Gradients were fractionated as described for polysome analysis, and monosome fractions were collected and reserved. Samples were adjusted to 0.5% SDS and digested with 200 µg ml<sup>-1</sup> proteinase K at 50 °C for 45 min. RNA was extracted using the hot acid phenol method, and precipitated with NaOAc and isopropanol. RNA was next resuspended in 500 µl 10 mM Tris 8 plus 2.5 µl SupersaseIn and centrifuged for 28 min in a Millipore YM100 micro-concentrator to remove RNA fragments longer than 100 nucleotides. RNA was precipitated from flow-through and separated on a 15% TBE-Urea gel, which revealed a clear band at ~30 nucleotides. RNA was extracted from this region of the gel and quantified using an Agilent BioAnalyzer. Samples were normalized to equivalent concentrations, and prepared for small RNA Illumina sequencing precisely as previously described<sup>8</sup>.

For total mRNA isolation, vehicle- or Torin-1-treated cells were washed in ice-cold PBS— and lysed in total RNA lysis buffer (15 mM HEPES-KOH (pH 7.4), 15 mM MgCl<sub>2</sub>, 0.3 M NaCl, 1.0% SDS). Lysates were homogenized by serial passage through a 21G needle and incubated at 65 °C for 5 min. RNA was then extracted using the hot acid phenol method and re-suspended in 200 µl 10 mM Tris 8. PolyA+ RNA was isolated from 150 µg total RNA using oligo dT<sub>25</sub> Dynabeads (Invitrogen) according to the manufacturer's instructions, resuspended in 20 µl 10 mM Tris 8, and then fragmented by the addition of 20 µl fragmentation buffer (2 mM EDTA, 100 mM NaCO<sub>3</sub> pH 9.2) and incubation at 95 °C for 20 min. Fragmented RNA was precipitated and separated on a 15% TBE-Urea gel. Short RNA fragments were then isolated from the 30-nucleotide region and quantified using an Agilent Bioanalyzer. Samples were normalized to equal concentrations and prepared for Illumina small RNA sequencing in parallel with footprint samples. Footprint libraries were prepared in two biological replicates, whereas total transcript libraries were prepared as single replicates.

**RNA sequence analyses.** Before alignment, footprint and total mRNA libraries were processed to remove cloning artefacts. Processed reads were then aligned to a database of mouse rRNA sequences using the Bowtie short-read alignment program to remove contaminating reads<sup>34</sup>. These rRNA reads represented 40–80% of

the footprint libraries, consistent with previous work<sup>8,35</sup>. The remaining reads were then aligned to the mm9 mouse genome. Reads that failed to align to genomic positions were re-aligned to a database of RefSeq gene sequences to capture those mapping to splice junctions. For both alignments, two mismatches were allowed in a 25-nucleotide 'seed' region, and reads were required to align to a single unique location. The resulting aligned libraries were then mapped to gene models described by the RefSeq annotation, which was downloaded from the UCSC genome browser website in March, 2010. Reads mapping to a unique genomic location, but to multiple transcripts, were counted equally for each transcript. Many mammalian genes have duplicate pseudogenes that exist throughout the genome, thereby causing reads that map to those genes to be discarded because they fail to map to a single unique location. To avoid under-counting these genes, the alignment process was repeated for previously unaligned reads using the same parameters, but allowing alignment at up to five unique sites. Counts for these 'multi-reads' were then distributed to each mapped gene model according to their relative representation in our library of unique alignments, similar to a strategy described previously<sup>36</sup>. Expression values were calculated as a modified version of RPKM, which normalizes mapped reads to gene length and library size<sup>36</sup>. The original RPKM value is calculated as  $R_i = 10^9 (C_i / L_i)$ , where  $C_i$  is the number of reads mapped to exons of gene  $i$ ,  $N$  is the number of mapped reads in the entire library, and  $L_i$  is the length of the spliced gene in nucleotides. Because contaminating rRNA constitutes a large part of our sequenced library, we calculated  $N$  as the number of reads mapping to exons of coding gene models. Translational efficiency was calculated as footprint (RPKM)/mRNA (RPKM). Values from biological replicates were averaged together.

Several additional constraints were applied before calculating changes in translational efficiency between vehicle- and Torin-1-treated conditions. First, we considered only genes where the combined number of reads between vehicle- and Torin-1-treated conditions exceeded 128. As reported previously, the replicate error in fold-change calculations for genes with fewer reads is primarily due to simple binomial sampling error<sup>8</sup>. Second, we considered only transcripts encoding protein-coding genes, and further excluded a small number of transcripts where greater than 25% of footprint reads mapped to introns. Three pseudogenes with homology to ribosomal protein coding genes were removed manually (NM\_001081036 and NM\_001111116, NM\_001101561). Last, we calculated translational changes for histone mRNAs from footprint reads only. Histone mRNAs were present in footprint libraries, but, because they have no polyA tail, could not be reliably detected in total mRNA libraries.

Analysis of footprint libraries can be used to determine the proportion of ribosomes engaged in translating each mRNA within each sample (RPM, reads per million), but can't directly measure differences in the overall number of translating ribosomes between samples. However, measurements of <sup>35</sup>S Cys/Met incorporation show that mTOR inhibition reduces the overall rate of protein synthesis by ~65%. Because mTOR primarily regulates initiation steps in translation, and because polysome analysis of Torin-1-treated cells clearly demonstrates a severe defect in initiation, the reduction in <sup>35</sup>S Cys/Met incorporation predominantly reflects a reduction in the number of translating ribosomes. Therefore, an mRNA that is translated by the same proportion of ribosomes in vehicle- and Torin-1-treated conditions is translated by approximately 65% fewer ribosomes when mTOR is inhibited. This factor can be incorporated as  $\Delta_{\text{translation}} = \Delta_{\text{ribosome density}} \times \Delta_{\text{translating ribosomes}}$  for each mRNA, where  $\Delta_{\text{ribosome density}}$  is the change in footprint RPM and  $\Delta_{\text{translating ribosomes}} = 0.35$  (65% reduction). Applying this correction, we find that all but four mRNAs are translated by fewer ribosomes in Torin-1-treated cells than in vehicle-treated cells.

**Complexity of 5' UTRs.** The 5' UTR sequences of all transcripts were obtained from the NCBI Reference Sequence collection. The minimum folding  $\Delta G^\circ$  were predicted for each sequence using QuikFold2 (<http://mfold.rna.albany.edu/?q=DINAMelt/Quickfold>) with the default parameters for version 3.0 of the RNA folding energy rules<sup>37</sup>.

**Gene ontology and TOP-like analysis.** To determine enrichment for gene ontology categories, mRNAs with a z-score greater than 1.5 or less than -1.5 were analysed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) website (<http://david.abcc.ncifcrf.gov/>)<sup>38,39</sup>. The 4,840 protein-coding mRNAs that were detected in both wild-type and DKO cells were used as a background reference set. Functional categories were clustered using the Functional Annotation Clustering tool, and representative gene ontology categories from each clustered set with a  $P$  value < 0.05 and FDR < 25% were selected.

For determination of TOP-like motifs, TSSs were first identified according to the experimentally determined database of TSSs from mouse NIH3T3 cells (dbTSS)<sup>19</sup>. For each mRNA, only promoters with a TSS tag of parts per million (p.p.m.) > 5 were considered. An mRNA was determined as TOP if the first nucleotide was a C followed by at least four pyrimidines, and TOP-like if it

contained a sequence of at least five pyrimidines within four nucleotides of either the most frequently detected TSS or a clear secondary TSS with a TSS-tag count of at least 30% of the primary site. Because dbTSS does not have confident TSS determinations for all mRNAs, and because data are from a single mouse cell line (3T3), we also considered TSS annotations from RefSeq, Ensembl or UCSC resources (downloaded from the UCSC genome browser May, 2011).

30. Liu, Q. *et al.* Discovery of 1-(4-(4-propionylpiperazin-1-yl)-3-(trifluoromethyl)phenyl)-9-(quinolin-3-yl)benzo[h][1,6]naphthyridin-2(1H)-one as a highly potent, selective mammalian target of rapamycin (mTOR) inhibitor for the treatment of cancer. *J. Med. Chem.* **53**, 7146–7155 (2010).
31. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298 (2006).
32. Ali, S. M. & Sabatini, D. M. Structure of S6 kinase 1 determines whether raptor-mTOR or rictor-mTOR phosphorylates its hydrophobic motif site. *J. Biol. Chem.* **280**, 19445–19448 (2005).
33. Levine, E. M., Becker, Y., Boone, C. W. & Eagle, H. Contact inhibition, macromolecular synthesis, and polyribosomes in cultured human diploid fibroblasts. *Proc. Natl Acad. Sci. USA* **53**, 350–356 (1965).
34. Langmead, B. Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinformatics* Chapter 11, Unit 11.7 (2010).
35. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–840 (2010).
36. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
37. Markham, N. R. & Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **33**, W577–W581 (2005).
38. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57 (2009).
39. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

# A *Xanthomonas* uridine 5'-monophosphate transferase inhibits plant immune kinases

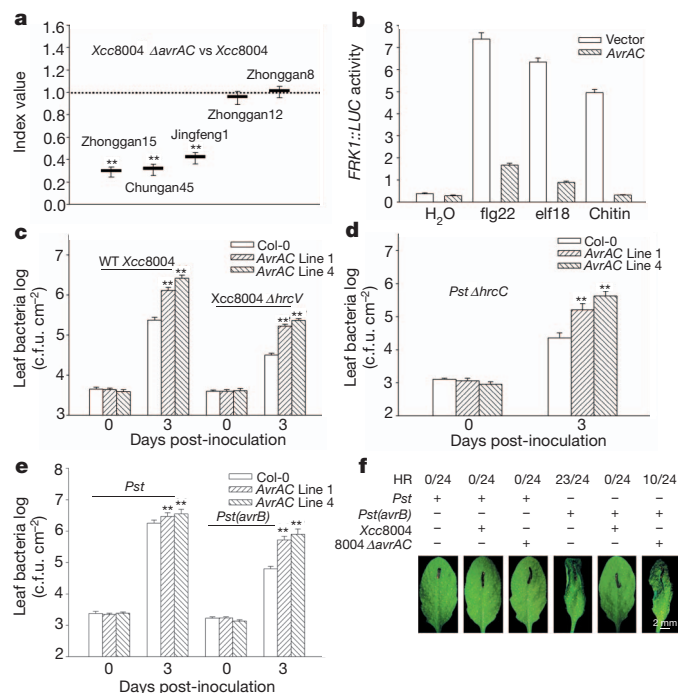
Feng Feng<sup>1,2,3</sup>, Fan Yang<sup>3</sup>, Wei Rong<sup>4</sup>, Xiaogang Wu<sup>3</sup>, Jie Zhang<sup>3</sup>, She Chen<sup>3</sup>, Chaozu He<sup>2,4</sup> & Jian-Min Zhou<sup>3,5</sup>

Plant innate immunity is activated on the detection of pathogen-associated molecular patterns (PAMPs) at the cell surface, or of pathogen effector proteins inside the plant cell<sup>1–4</sup>. Together, PAMP-triggered immunity and effector-triggered immunity constitute powerful defences against various phytopathogens. Pathogenic bacteria inject a variety of effector proteins into the host cell to assist infection or propagation. A number of effector proteins have been shown to inhibit plant immunity<sup>5</sup>, but the biochemical basis remains unknown for the vast majority of these effectors. Here we show that the *Xanthomonas campestris* pathovar *campestris* type III effector AvrAC enhances virulence and inhibits plant immunity by specifically targeting *Arabidopsis* BIK1 and RIPK, two receptor-like cytoplasmic kinases known to mediate immune signalling<sup>6–8</sup>. AvrAC is a uridylyl transferase that adds uridine 5'-monophosphate to and conceals conserved phosphorylation sites in the activation loop of BIK1 and RIPK, reducing their kinase activity and consequently inhibiting downstream signalling.

*Xanthomonas campestris* pathovar *campestris* (*Xcc*) is a causal agent of black rot diseases on numerous crucifer plants such as *Brassica* and *Arabidopsis*. AvrAC exists in all three sequenced strains of *Xcc* and a newly sequenced *Xanthomonas campestris* pathovar *raphani* strain, but not other *Xanthomonas* species. (<http://www.xanthomonas.org/t3e.html>). In *Xcc* strain 8004 (*Xcc*8004), AvrAC is recognized by *Arabidopsis* plants to trigger a weak effector-triggered immunity (ETI) in vascular, but not mesophyll tissues<sup>9</sup>. To determine if AvrAC has a role in virulence, we generated an *avrAC* knockout strain (*Xcc*8004 $\Delta$ *avrAC*) and performed competitive index assays on cabbage plants (Methods). The competitive index values (*Xcc*8004 $\Delta$ *avrAC* to *Xcc*8004 ratios) were 0.5–0.3 in 3/5 cultivars tested in multiple experiments (Fig. 1a), demonstrating a role of AvrAC in virulence on these plants. Well known PAMPs, including the bacterial flagellar peptide flg22, bacterial elongation factor (EF-Tu) peptide elf18, and fungal cell wall component chitin<sup>10</sup>, are recognized by pattern recognition receptors (PRRs) FLS2, EFR and CERK1, respectively<sup>4</sup>, to induce the expression of a PAMP-responsive reporter gene *FRK1::LUC*. The presence of the AvrAC transgene largely inhibited *FRK1::LUC* induction in protoplasts (Fig. 1b and Supplementary Fig. 1), indicating that AvrAC is a potent inhibitor of the PAMP-triggered immunity (PTI) signalling pathway.

Two AvrAC transgenic lines were generated to test if AvrAC impedes mesophyll resistance to *Xcc*8004, a nonpathogenic mutant bacteria *Xcc*8004 $\Delta$ *hrcV* and *Pseudomonas syringae* pathovar *tomato* (*Pst*)  $\Delta$ *hrcC* mutant bacteria. In all three cases, the transgenic lines supported 0.7–1.1 log more bacterial growth than did the wild-type plants (Fig. 1c, d). Thus AvrAC can markedly compromise PTI resistance to these bacteria. The AvrAC transgenic lines were also inoculated with *Pst* strains carrying effector genes *avrB*, *avrRpt2* or *avrPphB*, which elicit ETI resistance specified by the cytoplasmic immune receptors RPM1, RPS2 and RPS5, respectively<sup>3</sup>. The transgenic lines supported 0.9–1.0 log more *Pst*(*avrB*) bacterial growth than did wild-type

plants (Fig. 1e). In contrast, the transgenic lines supported less than 0.3 log increase in bacterial growth when *Pst*(*avrRpt2*), *Pst*(*avrPphB*) and *Pst* bacteria lacking an *avr* gene were used (Fig. 1e and Supplementary Fig. 2). Thus, although the AvrAC transgene compromised RPM1-dependent ETI, it did not appear to affect RPS2- and RPS5-dependent ETI. To determine further if bacterially delivered AvrAC similarly interferes with RPM1 activation, we infiltrated *Arabidopsis* plants with *Xcc*8004 before *Pst*(*avrB*) inoculation. Whereas the control plants inoculated with *Pst*(*avrB*) alone developed strong a hypersensitive



**Figure 1 | AvrAC inhibits plant immunity.** **a**, *avrAC* promotes *Xcc* bacterial virulence on cabbage plants. Competitive index value represents an average ratio of *Xcc*8004 $\Delta$ *avrAC* to *Xcc*8004 bacterial growth on the indicated cabbage cultivars (mean  $\pm$  s.d.;  $n \geq 6$ ;  $**P < 0.01$ , Student's *t*-test). **b**, Transient expression of AvrAC-Flag in protoplasts inhibits PAMP-induced *FRK1::LUC* transcription (mean  $\pm$  s.d.;  $n \geq 3$ ). **c–e**, The AvrAC transgene compromises PTI and RPM1-specified ETI in plants. Wild-type (WT; Col-0) and AvrAC transgenic lines pre-treated with oestradiol were infiltrated with *Xcc*8004 and *Xcc*8004 $\Delta$ *hrcV* (**c**), *Pst* $\Delta$ *hrcC* (**d**), or *Pst* or *Pst*(*avrB*) (**e**). Bacterial population in the leaf was measured at the indicated times (mean  $\pm$  s.d.;  $n \geq 6$ ;  $**P < 0.01$ , Student's *t*-test). **f**, Bacterially delivered AvrAC inhibits the RPM1-specified hypersensitive response (HR). Wild-type *Arabidopsis* plants were infiltrated with the indicated *Xcc* strains 1 day before infiltration of *Pst* or *Pst*(*avrB*), and development of the hypersensitive response was photographed 6 h later. The ratio of leaves showing a hypersensitive response to the total number of injected leaves is shown on top.

<sup>1</sup>School of Life Sciences, Tsinghua University, Beijing 100084, China. <sup>2</sup>Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China. <sup>3</sup>National Institute of Biological Sciences, Beijing 102206, China. <sup>4</sup>Hainan Key Laboratory for Sustainable Utilization of Tropical Bioresource, Hainan University, Haikou, Hainan 570228, China. <sup>5</sup>State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China.



response, the plants pre-treated with *Xcc8004* developed no hypersensitive response (Fig. 1f), indicating a strong interference of RPM1 activation by *Xcc8004*. The *Xcc8004**AvrAC* mutant strain was less capable of preventing hypersensitive response development (Fig. 1f and Supplementary Fig. 3), indicating that *AvrAC* interferes with RPM1-mediated ETI when delivered from *Xcc*.

To identify host targets and the mechanism by which *AvrAC* inhibits PTI, we systematically investigated various molecular events of PTI signalling pathways<sup>11</sup>. *flg22* perception triggers  $H_2O_2$  production and MAP kinase (MAPK) activation in wild-type plants within several minutes; both events were completely abolished or diminished in *AvrAC* transgenic plants (Fig. 2a, b). However, *AvrAC* was unable to inhibit MAPK activation triggered by the constitutively active MKK5<sup>DD</sup> (ref. 10; Supplementary Fig. 4). These results suggested that *AvrAC* acts upstream of MAPK cascades and  $H_2O_2$  production to inhibit PTI signalling.

PAMP detection rapidly induces phosphorylation of BIK1 (refs 6, 7), which acts immediately downstream of FLS2, EFR and CERK1 (ref. 7). The expression of *AvrAC* in protoplasts prevented the *flg22*-induced BIK1 phosphorylation, as indicated by a slow migrating form on SDS–polyacrylamide gel electrophoresis (SDS–PAGE; Fig. 2c). Instead, BIK1 uniformly migrated between the unphosphorylated and phosphorylated forms independent of *flg22* treatment.

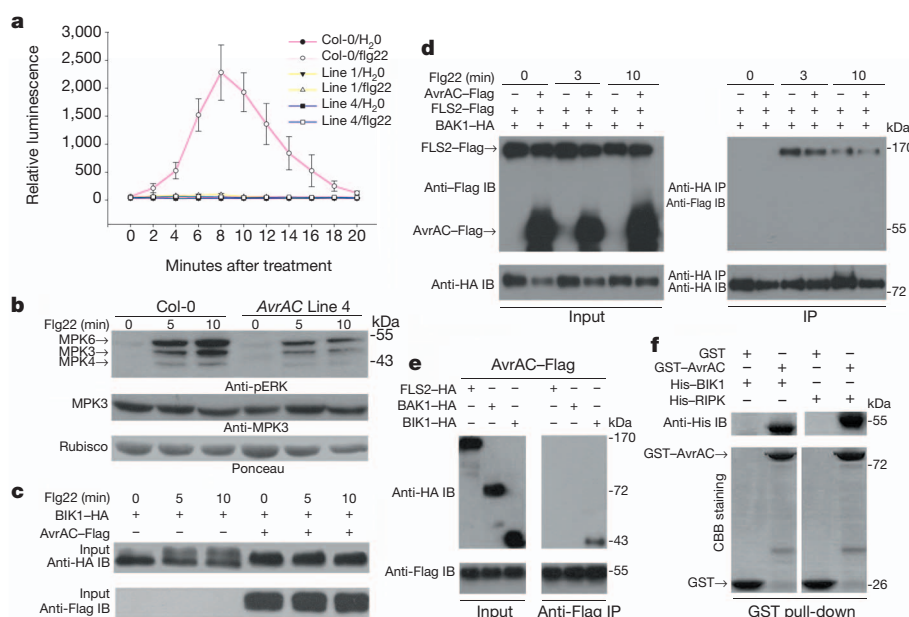
Upstream of BIK1 phosphorylation is the *flg22*-induced association between FLS2 and a receptor-like kinase called BAK1, forming an active receptor complex<sup>12,13</sup>. This interaction occurred normally in protoplasts irrespective of the presence or absence of *AvrAC* (Fig. 2d).

The results described above suggested that *AvrAC* directly target BIK1, FLS2 or BAK1. Co-immunoprecipitation assays showed that *AvrAC* can interact with BIK1, but not FLS2 or BAK1 (Fig. 2e). Glutathione *S*-transferase (GST) pull-down assays indicated that this interaction is direct (Fig. 2f). RIPK, an RLCK related to BIK1, can phosphorylate the RPM1-interacting protein RIN4 at T166 (ref. 8), and the phosphorylated RIN4 is responsible for triggering RPM1-specified ETI<sup>8,14</sup>. *AvrAC* also interacted with RIPK *in vitro* and in

protoplasts (Figs 2f and 4b). We conclude that BIK1 and RIPK are specific targets of *AvrAC*, a conclusion consistent with its inhibition of PTI and the RPM1-specific ETI.

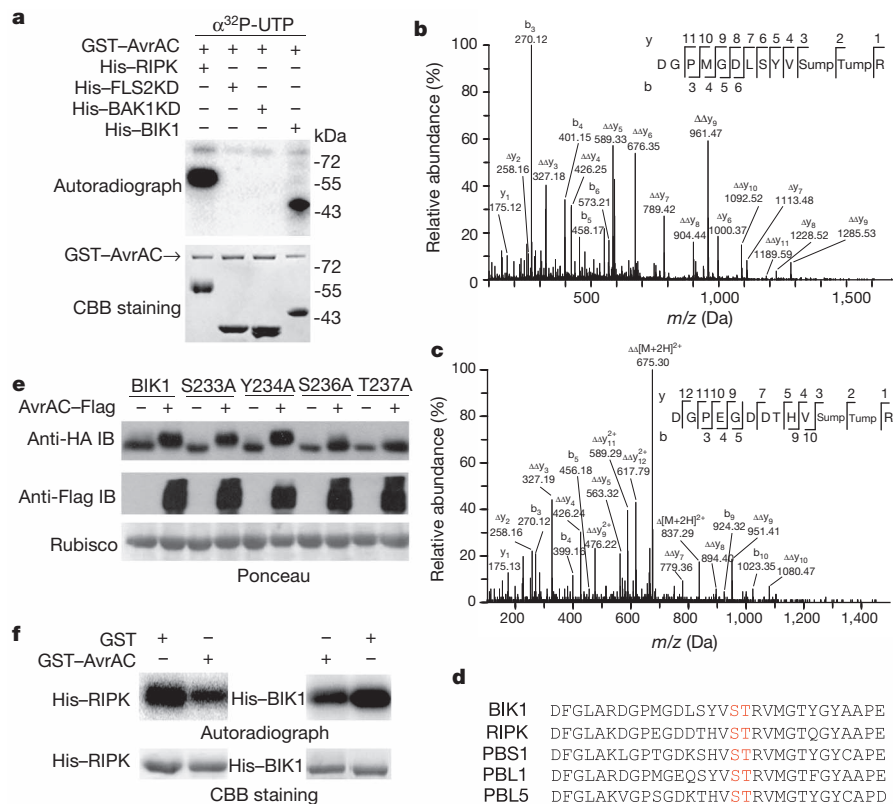
The *AvrAC*-induced electrophoresis mobility shift of BIK1 was resistant to phosphatase treatment (Supplementary Fig. 5a). Co-expression of BIK1 and *AvrAC* in *Escherichia coli* detected the same phosphatase-resistant, slow migrating form of BIK1 (Supplementary Fig. 5b, c), suggesting an *AvrAC*-induced BIK1 modification that is not caused by protein phosphorylation.

*AvrAC* contains a filamentation-induced by c-AMP domain (Fic) in the carboxy terminus. More than 4,300 Fic-domain-containing proteins have been found in bacteria and animals, mostly with unknown biochemical function<sup>15,16</sup>. Two Fic-domain-containing, animal bacterial pathogen effectors *Vibrio* VopS and *Histophilus* IbpA catalyse adenylation (AMPylation) on threonine or tyrosine residues of Rho family GTPases to disrupt cytoskeleton in the host cell<sup>17,18</sup>. Another bacterial effector protein, *Legionella pneumophila* DrrA (also known as SidM) uses a glutamine synthetase adenylyl transferase domain to adenylylate another host GTPase called Rab1b<sup>19</sup>. However, *in vitro* labelling and mass spectrometric assays with *AvrAC* and BIK1 recombinant proteins failed to detect any AMPylation on BIK1. Surprisingly, mass spectrometry of the *AvrAC* recombinant protein (expressed alone in *E. coli*) revealed multiple modified protein components with serial additions of 306 Da, indicative of uridine 5'-monophosphate (UMP) modifications (Supplementary Fig. 6a). Tandem mass spectrometry (MS/MS) revealed multiple serine and threonine residues in *AvrAC* containing mono-UMP modifications (Supplementary Fig. 6b). The results suggested that *AvrAC* is a UMP transferase. *In vitro* assays with  $\alpha^{32}P$ -UTP showed that both BIK1 and RIPK were uridylylated only when incubated with *AvrAC* (Fig. 3a). *AvrAC* also labelled itself in this assay, explaining the UMP modifications on *AvrAC* described above. PBL1, a homologue of BIK1 that acts additively with BIK1 in PTI signalling<sup>7</sup>, was also modified by *AvrAC* (Supplementary Fig. 7). Interestingly, another RLCK VII subfamily member, PBS1, was not modified under



**Figure 2 | *AvrAC* targets BIK1 and inhibits early PTI signalling events.** **a**, *Flg22*-induced  $H_2O_2$  production in wild-type and *AvrAC* transgenic plants (mean  $\pm$  s.d.;  $n \geq 8$ ). **b**, *Flg22*-induced MAPK phosphorylation in wild-type and *AvrAC* transgenic plants. The identities of phosphorylated MAPKs are labelled as described (Methods). Equal loading is indicated by immunoblot with anti-MPK3 antibodies or ponceau staining of Rubisco protein. **c**, *Flg22*-induced BIK1-HA phosphorylation in protoplasts in the presence or absence of *AvrAC*-Flag. **d**, *Flg22*-induced FLS2–BAK1 association in protoplasts in the

presence of *AvrAC* as determined by co-immunoprecipitation assays. The minor differences of BAK1–HA levels resulted from sample processing. **e**, *AvrAC* interacts with BIK1, but not FLS2 and BAK1, in protoplasts. BIK1–HA, BAK1–HA and FLS2–HA were co-expressed with *AvrAC*-Flag in protoplasts and subjected to co-immunoprecipitation. **f**, *AvrAC* interacts with BIK1 and RIPK *in vitro* in a GST pull-down assay. The bound proteins were analysed by anti-His immunoblot and the membrane stained with Coomassie blue (CBB). IB, immunoblot; IP, immunoprecipitation.



**Figure 3 | AvrAC uridylylates BIK1 and RIPK.** **a**, AvrAC uridylylates BIK1 and RIPK, but not FLS2 and BAK1, *in vitro*. GST-AvrAC was incubated with His-BIK1, His-RIPK, or His-tagged kinase domains (KD) of FLS2 or BAK1, in the presence of  $\alpha^{32}\text{P}$ -UTP. The resulting products were subjected to SDS-PAGE and autoradiography. **b**, **c**, AvrAC uridylylates conserved serine and threonine residues in the activation loop of BIK1 and RIPK in the plant cell. BIK1-Flag (**b**) or RIPK-Flag (**c**) was co-expressed with AvrAC-HA in protoplasts, affinity purified, and subjected to MS/MS spectrometric analyses. The b- and y- ions are marked in the spectrum and illustrated along the peptide sequence above the spectrum. This unambiguously identified BIK1 Ser 236 and Thr 237, and RIPK Ser 251 and Thr 252 as UMP-modified residues.  $\Delta$

the same condition. The same assay failed to label the kinase domain of FLS2 and BAK1, indicating that certain members of the RLCK VII subfamily are specific substrates of AvrAC (Fig. 3a). MS/MS analysis on the recombinant BIK1 protein that had been co-expressed with AvrAC uncovered a tryptic peptide (DGPMGDLSTYVSTR) derived from the activation loop containing UMP at Ser 236 and Thr 237 (Supplementary Fig. 8a). Likewise, RIPK that had been co-expressed with AvrAC also produced a tryptic peptide (DGPEGDDTHVSTR) located in the activation loop in which Ser 251 and Thr 252 were UMP modified (Supplementary Fig. 8b). BIK1 and RIPK proteins that had been co-expressed with AvrAC in protoplasts also contained the same UMP modifications (Fig. 3b, c), indicating that AvrAC can uridylylate BIK1 and RIPK in the plant cell. The vast majority of the BIK1 and RIPK activation loop peptides were uridylylated (Supplementary Fig. 9). Together these results unequivocally demonstrated that AvrAC is a potent enzyme that uridylylates BIK1 and RIPK both *in vitro* and in plant cells.

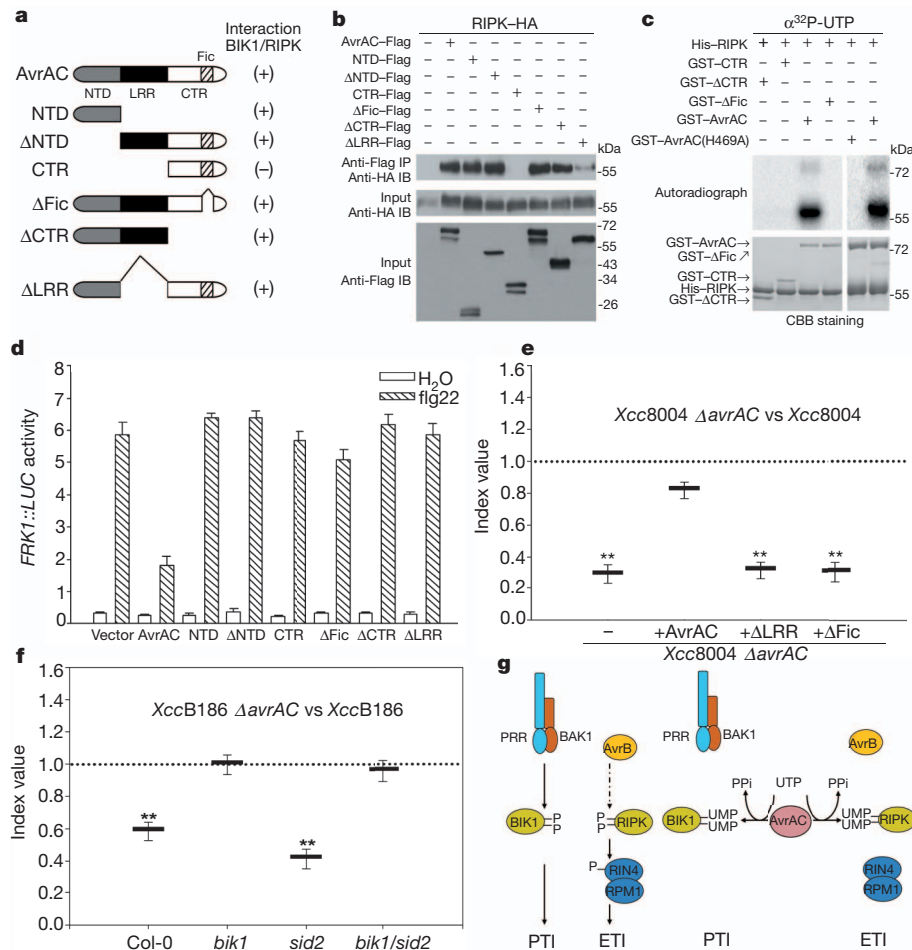
ATP is the preferred nucleotide substrate for IbpA and VopS to modify Rho GTPases *in vitro*<sup>16</sup>, although GTP, UTP and CTP can also be used. Whether nucleotides other than ATP are used by these effectors to modify Rho GTPases in host cells remains unknown. AvrAC failed to AMPylate BIK1, and we observed no AMPylation on any peptides of BIK1, RIPK or AvrAC expressed in *E. coli* or plant cells. Another Fic-domain protein, AnkX, from *Legionella pneumophila* was recently shown to covalently attach a phosphocholine moiety to the Rho GTPase Rab1 (ref. 20). The commonality shared by the three

represents neutral loss of an UMP group with nominal mass of 324 Da. **d**, Alignment of the activation loop of selected RLCK VII kinases. The UMP modification sites are highlighted in red. **e**, BIK1 Ser 236 and Thr 237 are required for uridylylation by AvrAC. Various BIK1-HA mutants carrying the indicated amino acid substitutions were co-expressed with AvrAC-Flag and analysed by immunoblot. Note the increased BIK1 protein level in the presence of AvrAC, suggesting an AvrAC-induced stabilization of BIK1 protein. **f**, Modification by AvrAC inhibits BIK1 and RIPK kinase activity. His-BIK1 or His-RIPK co-expressed with GST-AvrAC or GST in *E. coli* was affinity purified and subjected to autophosphorylation assay.

modifications suggests that Fic-domain proteins are phosphotransferases with different substrates.

BIK1 Ser 236 and Thr 237 and RIPK Ser 251 and Thr 252 are highly conserved among RLCK VII subfamily members (Fig. 3d). Phosphorylation of BIK1 Ser 236 and Thr 237 has a crucial role in PTI signalling<sup>6,7,21</sup>. AvrAC failed to induce the BIK1(S236A) and BIK1(T237A) mutant protein migration in protoplasts (Fig. 3e), indicating that these residues are required for UMP modification. A similar result was obtained with a BIK1(S233A,Y234A,S236A,T237A) quadruple mutant<sup>7</sup> (Supplementary Fig. 10). In contrast, mutations in Lys 105, an invariant residue required for ATP binding, and Ser 233 and Tyr 234, two less conserved amino acid residues in the activation loop, did not affect the AvrAC-induced migration (Supplementary Fig. 10 and Fig. 3e). RIPK and BIK1 that had been UMP modified by AvrAC in *E. coli* displayed reduced autophosphorylation (Fig. 3f). Taken together, these results demonstrated that AvrAC inhibits BIK1 and RIPK kinase activities by directly uridylylating the conserved phosphorylation sites in the activation loop, thereby preventing their phosphorylation and subsequent signal transduction.

AvrAC contains an N-terminal domain (NTD) of unknown function, a leucine-rich-repeat domain (LRR), and a C-terminal region (CTR) harbouring the Fic domain (Fig. 4a). The homology of the AvrAC Fic domain with other Fic domains mainly resides in the C-terminal subdomain (Supplementary Fig. 11), in which the highly conserved HPFx(D/E)GN(G/K)R motif is responsible for nucleotide binding and catalysis<sup>22,23</sup>. In particular, the invariant His residue



**Figure 4 | AvrAC uridylyl transferase activity and BIK1 are required for AvrAC virulence in plants.** **a**, Schematic diagram of AvrAC mutant constructs and their interaction with BIK1 and RIPK. The ability of these constructs to interact with BIK1 and RIPK is summarized on the right. **b**, The AvrAC NTD and LRR, but not CTR, interacted with RIPK in co-immunoprecipitation assays. **c**, *In vitro* uridylylation of RIPK by the indicated GST-AvrAC recombinant proteins. **d**, Inhibition of flg22-induced *FRK1::LUC* expression by AvrAC mutants (mean  $\pm$  s.d.;  $n \geq 3$ ). **e**, The LRR domain and Fic motif are required for AvrAC virulence in cabbage plants. Competitive index assays were carried out on cabbage cultivar Zhonggan15 with *Xcc8004*  $\Delta$ avrAC complemented with the indicated constructs (mean  $\pm$  s.d.;  $n \geq 6$ ; \*\* $P \leq 0.01$ , Student's *t*-test). **f**, AvrAC

corresponding to AvrAC His 469 has a key role in the catalysis. AvrAC constructs lacking the Fic catalytic motif ( $\Delta$ Fic) or CTR ( $\Delta$ CTR) were fully capable of interacting with BIK1 and RIPK, whereas the AvrAC CTR alone was completely unable to interact with BIK1 and RIPK (Fig. 4b and Supplementary Fig. 12). Furthermore, the AvrAC NTD alone, the construct lacking LRR, and the construct lacking NTD or CTR, or substitution of His 469 with Ala, rendered AvrAC completely inactive in RIPK uridylylation *in vitro* (Fig. 4c), and induced a BIK1 mobility shift in protoplasts (Supplementary Figs 12 and 13). These results are consistent with the notion that the AvrAC Fic domain is responsible for catalysis, whereas NTD and LRR allow docking of substrate proteins.

We next determined the role of individual AvrAC domains in the inhibition of PTI signalling and virulence function. All the AvrAC deletion constructs were unable to inhibit *FRK1::LUC* expression, whereas the AvrAC(H469A) mutant was partially affected (Fig. 4d, Supplementary Figs 14 and 15). Only the wild-type, but not the truncated versions of *avrAC*, was able to promote *Xcc8004*  $\Delta$ avrAC bacterial growth on Zhonggan15 plants (Fig. 4e). Together these

results indicated that the ability of AvrAC to inhibit immune signalling and enhance virulence not only required its binding to, but also UMP modification of, the substrate proteins.

*Xcc8004* and *Xcc8004*  $\Delta$ avrAC are weakly virulent on *Arabidopsis* plants<sup>9</sup>, preventing a test of AvrAC virulence function in *Arabidopsis*. We found that *Xcc* strain B186, which is highly virulent on *Arabidopsis* plants<sup>24</sup>, contains AvrAC with a single amino acid substitution (AvrAC(T358A)). We then generated an *XccB186*  $\Delta$ avrAC mutant and conducted competitive index assays in *Arabidopsis* mesophyll tissues. *XccB186*  $\Delta$ avrAC grew to ~60% of *XccB186* in wild-type *Arabidopsis* plants, and the difference was statistically significant (Fig. 4f). In contrast, the two strains grew to nearly identical levels in *bik1* mutant plants. The *bik1* mutant is known to accumulate high levels of the defence hormone salicylate<sup>25</sup>, which may potentially affect the virulence assay. We therefore carried out the assays on the *sid2* mutant, which has greatly reduced salicylate production. *XccB186*  $\Delta$ avrAC grew to ~40% of *XccB186* on *sid2* plants (Fig. 4f). In contrast, the two strains grew to nearly identical levels in the *bik1/sid2* double-mutant plants, again indicating that *avrAC* promotes *XccB186* bacterial growth in a BIK1-dependent manner. Together these results demonstrated that BIK1 is a key target for AvrAC virulence activity.

results indicated that the ability of AvrAC to inhibit immune signalling and enhance virulence not only required its binding to, but also UMP modification of, the substrate proteins.

*Xcc8004* and *Xcc8004*  $\Delta$ avrAC are weakly virulent on *Arabidopsis* plants<sup>9</sup>, preventing a test of AvrAC virulence function in *Arabidopsis*. We found that *Xcc* strain B186, which is highly virulent on *Arabidopsis* plants<sup>24</sup>, contains AvrAC with a single amino acid substitution (AvrAC(T358A)). We then generated an *XccB186*  $\Delta$ avrAC mutant and conducted competitive index assays in *Arabidopsis* mesophyll tissues. *XccB186*  $\Delta$ avrAC grew to ~60% of *XccB186* in wild-type *Arabidopsis* plants, and the difference was statistically significant (Fig. 4f). In contrast, the two strains grew to nearly identical levels in *bik1* mutant plants. The *bik1* mutant is known to accumulate high levels of the defence hormone salicylate<sup>25</sup>, which may potentially affect the virulence assay. We therefore carried out the assays on the *sid2* mutant, which has greatly reduced salicylate production. *XccB186*  $\Delta$ avrAC grew to ~40% of *XccB186* on *sid2* plants (Fig. 4f). In contrast, the two strains grew to nearly identical levels in the *bik1/sid2* double-mutant plants, again indicating that *avrAC* promotes *XccB186* bacterial growth in a BIK1-dependent manner. Together these results demonstrated that BIK1 is a key target for AvrAC virulence activity.



In this study we demonstrated that AvrAC inhibits both PTI and RPM1-specified ETI by targeting two related RLCKs, BIK1 and RIPK, as illustrated in the model (Fig. 4g). BIK1 and PBL1 are not required for PAMP-induced MAPK activation (Supplementary Fig. 16), suggesting that additional host targets account for the MAPK inhibition by AvrAC. Although *Xcc* is not known to trigger the RPM1-specific ETI in *Arabidopsis*, the *Xcc* bacterium carries XopAH (also known as AvrXccC), which is highly homologous to AvrB<sup>26</sup>. It is formally possible that XopAH triggers ETI specified by an unknown resistance protein in some plants through RIPK. Inactivation of RIPK by AvrAC may allow the *Xcc* bacterium to inhibit such a potential ETI resistance in plants. AvrAC is the only Fic-domain effector among all sequenced phytopathogenic bacteria. However, the Fic and Doc (death on curing) domains contain a core structure, designated Fido, which is shared by AvrB and XopAH<sup>27</sup>. AvrB and XopAH do not carry the catalytic His residue conserved in Fic and Doc proteins, and it remains to be determined whether they also possess phosphotransferase activities.

AvrAC uridylylates and masks the conserved phosphorylation sites in the activation loop of BIK1 and RIPK, preventing the activation of these kinases and impairing plant immunity. Interestingly, a similar strategy has been used by the *Yersinia* effector YopJ, which selectively acetylates serine and threonine residues in the activation loop of human MEK6, thereby blocking the phosphorylation of the target protein and signalling<sup>28</sup>. Multiple *P. syringae* type III effectors are known to target plant protein kinases and inhibit PTI signalling by using diverse biochemical mechanisms<sup>5</sup>. The work described here illustrates a unique biochemical mechanism by which the *Xcc* bacterium combats the plant innate immune system.

## METHODS SUMMARY

**In vitro uridylylation assay.** Affinity-purified recombinant His-RIPK or His-BIK1 protein was incubated with GST-AvrAC in an uridylylation reaction buffer supplied with  $\alpha$ -<sup>32</sup>P UTP. The products were separated on a SDS-PAGE gel and visualized by autoradiography.

**Mapping UMP modification sites.** Affinity-purified GST-AvrAC, His-BIK1, His-RIPK, BIK1-Flag and RIPK-Flag proteins were separated by SDS-PAGE. The corresponding protein bands were sliced and in-gel digested with trypsin overnight at 37 °C. The resulting peptides were extracted and analysed by nano-flow LC-MS/MS on a QSTAR XL mass spectrometer. Database searches were performed on a Mascot server with the following variable modifications: oxidation on methionine, carbamidomethylation on cysteine, UMP modification on serine, threonine or tyrosine, and AMP modification on serine, threonine or tyrosine.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 5 September 2011; accepted 17 February 2012.

Published online 15 April 2012.

1. Ausubel, F. M. Are innate immune signaling pathways in plants and animals conserved? *Nature Immunol.* **6**, 973–979 (2005).
2. Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**, 803–814 (2006).
3. Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
4. Boller, T. & Felix, G. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).
5. Block, A. & Alfano, J. R. Plant targets for *Pseudomonas syringae* type III effectors: virulence targets or guarded decoys? *Curr. Opin. Microbiol.* **14**, 39–46 (2011).
6. Lu, D. *et al.* A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proc. Natl Acad. Sci. USA* **107**, 496–501 (2010).

7. Zhang, J. *et al.* Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe* **7**, 290–301 (2010).
8. Liu, J., Elmore, J. M., Lin, Z.-J. D. & Coaker, G. A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. *Cell Host Microbe* **9**, 137–146 (2011).
9. Xu, R.-Q. *et al.* AvrAC<sub>Xcc8004</sub>, a type III effector with a leucine-rich-repeat domain from *Xanthomonas campestris* pathovar *campestris* confers avirulence in vascular tissues of *Arabidopsis thaliana* ecotype Col-0. *J. Bacteriol.* **190**, 343–355 (2008).
10. He, P. *et al.* Specific bacterial suppressors of MAMP signaling upstream of MAPKKK in *Arabidopsis* innate immunity. *Cell* **125**, 563–575 (2006).
11. Segonzac, C. & Zipfel, C. Activation of plant pattern-recognition receptors by bacteria. *Curr. Opin. Microbiol.* **14**, 54–61 (2011).
12. Chinchilla, D. *et al.* A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**, 497–500 (2007).
13. Heese, A. *et al.* The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc. Natl Acad. Sci. USA* **104**, 12217–12222 (2007).
14. Chung, E.-H. *et al.* Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell Host Microbe* **9**, 125–136 (2011).
15. Woolery, A. R., Luong, P., Broberg, C. A. & Orth, K. AMPylation: something old is new again. *Front. Microbiol.* **1**, 113 (2010).
16. Mattoo, S. *et al.* Comparative analysis of *Histophilus somni* IbpA with other FIC enzymes reveals differences in substrate and nucleotide specificities. *J. Biol. Chem.* **286**, 32834–32842 (2011).
17. Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* **323**, 269–272 (2009).
18. Worby, C. A. *et al.* The Fic domain: regulation of cell signaling by adenylation. *Mol. Cell* **34**, 93–103 (2009).
19. Muller, M. P. *et al.* The *Legionella* effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science* **329**, 946–949 (2010).
20. Mukherjee, S. *et al.* Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103–106 (2011).
21. Laluk, K. *et al.* Biochemical and genetic requirements for function of the immune response regulator BOTRYTIS-INDUCED KINASE1 in plant growth, ethylene signaling, and PAMP-triggered immunity in *Arabidopsis*. *Plant Cell* **23**, 2831–2849 (2011).
22. Luong, P. *et al.* Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J. Biol. Chem.* **285**, 20155–20163 (2010).
23. Xiao, J. *et al.* Structural basis of Fic mediated adenylation. *Nature Struct. Mol. Biol.* **17**, 1004–1010 (2010).
24. Sun, W., Dunning, M. F., Pfund, C., Weingarten, R. & Bent, A. F. Within-species flagellin polymorphism in *Xanthomonas campestris* pv *campestris* and its impact on elicitation of *Arabidopsis* FLAGELLIN SENSING2-dependent defenses. *Plant Cell* **18**, 764–779 (2006).
25. Veronese, P. *et al.* The membrane-anchored BOTRYTIS-INDUCED KINASE1 plays distinct roles in *Arabidopsis* resistance to necrotrophic and biotrophic pathogens. *Plant Cell* **18**, 257–273 (2006).
26. White, F. F., Potnis, N., Jones, J. B. & Koebe, R. The type III effectors of *Xanthomonas*. *Mol. Plant Pathol.* **10**, 749–766 (2009).
27. Kinch, L. N., Yarbrough, M. L., Orth, K. & Grishin, N. V. Fido, a novel AMPylation domain common to Fic, Doc, and AvrB. *PLoS ONE* **4**, e5818 (2009).
28. Mukherjee, S. *et al.* *Yersinia* YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science* **312**, 1211–1214 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors thank J. Chai for sharing plasmids before publication, A. Bent for the XccB186 strain, S. Y. He and F. White for helpful comments. J.-M.Z. was supported by grants from the Chinese Ministry of Science and Technology (2011CB100700; 2010CB835301). S.C. was supported by a grant from the Chinese Ministry of Science and Technology (2010CB835204). C.H. was supported by Funds from Hainan University.

**Author Contributions** S.C., C.H. and J.-M.Z. conceived and designed the experiments. F.F., F.Y., W.R., X.W. and J.Z. performed the experiments. F.F., F.Y., S.C., C.H. and J.-M.Z. analysed the data. F.F., S.C., C.H. and J.-M.Z. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.-M.Z. ([zhoujianmin@nibs.ac.cn](mailto:zhoujianmin@nibs.ac.cn)) or C.H. ([czhe@hainu.edu.cn](mailto:czhe@hainu.edu.cn)).

## METHODS

**Plants and bacterial strains.** *Arabidopsis* plants were grown in a growth room at 23 °C and 70% relative humidity with a 10/14 h day/night light cycle for 5 weeks before protoplast isolation or bacterial inoculation. Cabbage cultivars Zhonggan15, Chunggan45, Jingfeng1, Zhonggan12, Zhonggan8 were grown in greenhouses at 25 °C. The bacterial strains used in this study included *Xcc8004* (ref. 9), *Xcc8004ΔhrcV* (ref. 29), *Xcc* strain B186 (*XccB186*) (ref. 23), *Pst* strain DC3000, *PstΔhrcC* (ref. 30), *Pst(avrB)*, *Pst(avrRpt2)*, *Pst(avrPphB)* and *Agrobacterium* GV3101.

**Constructs, generation of knockout strains and transgenic plants.** To generate recombinant protein constructs, full-length *AvrAC* and truncated mutants were PCR-amplified from *Xcc8004* genomic DNA and inserted between *EcoRI* and *XhoI* sites of pGEX-6p-1 (Amersham Biosciences) to generate GST-*AvrAC*, GST-CTR (C-terminal region), GST-ΔCTR and GST-ΔFic. A GST-*AvrAC*(H469A) mutant construct was generated by site-directed mutagenesis. The coding sequences of *Arabidopsis* *RIPK*, *BIK1* and *PBL1* were amplified from cDNA and inserted into pET28a or pET28b (Novagen) to generate His-*RIPK*, His-*BIK1* and His-*PBL1*.

To generate constructs for protoplast transfection assay, *RIPK*, *AvrAC* and its derivatives were PCR-amplified and inserted between *XhoI* and *Csp45I* sites of pUC19-35S-Flag-RBS vector<sup>31</sup> to generate *RIPK*-Flag, *AvrAC*-Flag, NTD-Flag, ΔNTD-Flag, CTR-Flag, ΔCTR-Flag, ΔFic-Flag, ΔLRR-Flag and *AvrAC*(H469A)-Flag. *RIPK* and *AvrAC* were cloned into *KpnI* and *Sall* sites of pUC19-35S-HA-RBS<sup>31</sup> to generate *RIPK*-HA and *AvrAC*-HA, respectively. *BIK1*-Flag, *BIK1*-HA and *BIK1* mutants were described previously<sup>7</sup>.

To generate *avrAC* knockout mutants in *Xcc8004* and *XccB186*, an internal sequence corresponding to nucleotides 30–500 of the *avrAC* coding region was PCR-amplified from genomic DNA of *Xcc8004* and cloned into the suicide plasmid pK18mob<sup>32</sup>. The recombinant plasmid was transferred into the wild-type *Xcc* strains by triparental conjugation, and the single crossover mutants were confirmed by PCR and designated *Xcc8004ΔavrAC* and *XccB186ΔavrAC*. To complement the *Xcc8004ΔavrAC* mutant, the 400 bp upstream sequence of the *avrAC* coding region was PCR-amplified from the genomic DNA of *Xcc8004* and ligated between the *XbaI* and *HindIII* sites of pLAFR6 (ref. 9), generating an intermediate plasmid carrying *avrAC* promoter sequence. A full-length *avrAC*, *avrACΔLRR* or *avrACΔFic* fragments were PCR-amplified using the pUC19-35S-Flag-RBS plasmids containing the corresponding *avrAC* fragments as templates (see above) before integrated into the intermediate plasmid. The resulting constructs were introduced into the *Xcc8004ΔavrAC* strain by triparental conjugation.

The *avrAC* coding region was PCR-amplified from *Xcc8004* genomic DNA, ligated into a modified pER8 vector<sup>31,33</sup>. The resulting clone containing *AvrAC*-*Flag* under the control of the oestrogen-inducible promoter was transformed into *Arabidopsis* (Col-0) by *Agrobacterium*-mediated transformation according to standard protocols. Transgenic plants were selected on Murashige and Skoog (MS) plates containing hygromycin. Two independent transgenic lines were selected for experiments. The transgenic plants were sprayed with 50 mM oestradiol in a 0.01% silwet L-77 solution for 12 h to induce *AvrAC* protein expression.

**Dual reporter assay.** *AvrAC* and its truncation mutants in the pUC19-35S-Flag-RBS vector were co-transfected with *FRK1::LUC* (firefly luciferase) and 35S:*RLUC* (Renilla luciferase) into *Arabidopsis* protoplasts. The protoplasts were incubated overnight under low light, treated with 1 μM flg22, 1 μM elf18, or 200 μg ml<sup>-1</sup> chitin (Sigma) for 3 h. Protein was then isolated, and *LUC* activity was recorded by using the Dual-Luciferase Reporter system (Promega) according to the manufacturer's instructions.

**Oxidative burst.** Oestradiol-induced leaves were sliced into 1 mm strips, incubated in 200 μl water in a 96-well plate for 12 h before the addition of 1 μM flg22 in 200 μl reaction buffer supplement with 20 mM luminol and 1 μg horseradish peroxidase (Sigma). Luminescence was recorded with a Luminometer (Promega) for 20 min.

**MAPK activity assay.** Two-week-old wild-type and transgenic seedlings grown on 1/2 MS medium were sprayed with 10 μM flg22 or water containing 0.02% Silwet L-77 for 10 min and frozen in liquid nitrogen. The seedlings were homogenized in an extraction buffer containing 50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 1 mM EDTA, 0.2% Triton-X 100, 1 mM DTT, complete protease inhibitors (Roche), and phosphatase inhibitors (Fisher). The protein concentration was determined using a Bio-Rad Bradford protein assay kit, and equal amounts of total protein were electrophoresed on 10% SDS-PAGE. An anti-PERK antibody (no. 4370S, Cell Signaling) was used to determine phosphorylation state of MPK3, MPK4 and MPK6 in an immunoblot. A duplicate blot was reacted with anti-MPK3 antibodies (Sigma) to determine the amount of total MPK3. For MAPK activation by MKK5<sup>DD</sup>, protoplasts prepared from wild-type plants were transfected with the indicated plasmids, and protein was isolated 12 h later for

immunoblot analyses. The identities of phosphorylated MAPKs were labelled as described previously<sup>34,35</sup>.

**Bacterial growth and hypersensitive response assays.** Five-week-old *Arabidopsis* leaves were infiltrated with the indicated *Xcc* or *Pst* bacteria at 10<sup>6</sup> c.f.u. ml<sup>-1</sup>, and the bacterial population in the leaf was counted at the indicated times. Each data point consisted of at least four replicates. For the hypersensitive response assay, wild-type plants were first infiltrated with *Xcc8004* or *Xcc8004ΔavrAC* mutant bacteria at 10<sup>7</sup> c.f.u. ml<sup>-1</sup> 1 day before infiltrating 5 × 10<sup>7</sup> c.f.u. ml<sup>-1</sup> *Pst* or *Pst(avrB)*. Plants were photographed 6 h after the second inoculation.

A competitive index assay<sup>36</sup> was used to compare virulence between wild-type and *ΔavrAC* mutant strains of *Xcc*. Briefly, equal amounts (c.f.u.) of wild-type and *ΔavrAC* mutant *Xcc* bacteria were mixed and inoculated into plants, and leaf bacterial number for each strain was determined by plating on NYG agar medium supplemented with appropriate antibiotics. The ratio between the mutant and wild-type bacteria (competitive index) was calculated at the desired times. For virulence assay on cabbage, fully expanded leaves of 4-week-old plants were clipping inoculated<sup>37</sup> with a mixture of *Xcc8004* and *Xcc8004ΔavrAC* bacteria at an equal concentration of 10<sup>8</sup> c.f.u. ml<sup>-1</sup>, maintained at high humidity for 48 h under a transparent dome. Leaf bacterial number and competitive index were determined 6 days after inoculation. For virulence assay on *Arabidopsis* plants, 5-week-old *Arabidopsis* leaves were infiltrated with a mixture of wild-type *XccB186* and *XccB186ΔavrAC* mutant bacteria at an equal concentration of 10<sup>6</sup> c.f.u. ml<sup>-1</sup>. Leaf bacterial number and competitive index were determined 3 days after inoculation.

**CyaA secretion assay.** The plasmid pCPP3221 containing an *AvrPto*-Cya fusion sequence was introduced into the wild-type *Pst* (DC3000) and its *ΔhrcC* mutant derivative. The derived strains were inoculated at 1 × 10<sup>8</sup> c.f.u. ml<sup>-1</sup> into *Arabidopsis* leaves that had been pre-treated with 10<sup>7</sup> c.f.u. ml<sup>-1</sup> *Xcc8004* or *Xcc8004ΔavrAC* for 24 h. Amounts of *AvrPto*-Cya secretion were determined as described previously<sup>38</sup>. Briefly, leaf samples were ground in liquid nitrogen and resuspended in 0.1 M HCl. The total protein content of each sample was determined using the Bradford assay (Bio-Rad). A direct cAMP (cAMP) immunoassay kit (Sigma) was used to measure cAMP levels following the manufacturer's instruction.

**Recombinant proteins expression and GST pull-down assay.** The recombinant proteins were affinity purified following the manufacturer's instruction. For pull-down assays, 10 μg each of His-*RIPK*, His-*BIK1* and GST-*AvrAC* were incubated on a rotator with glutathione agarose beads (GE Healthcare) in a buffer containing 25 mM Tris-HCl (pH 7.5), 100 mM NaCl and 1 mM DTT for 3 h, washed five times with a buffer containing 25 mM Tris-HCl (pH 7.5), 100 mM NaCl and 1 mM DTT. The bound protein was eluted with 15 mM GSH and immunoblotted with a mouse anti-His monoclonal antibody (Sigma-Aldrich).

**Co-immunoprecipitation assay.** The protoplasts were transfected with the indicated constructs, incubated for 12 h, treated with either H<sub>2</sub>O or 1 μM flg22 for 3–10 min, and total protein was isolated with an extraction buffer (see MAPK activity assay). For anti-Flag immunoprecipitation, total protein was incubated with an agarose-conjugated anti-Flag antibody (Sigma-Aldrich) for 4 h, washed six times with a buffer containing 50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 1 mM EDTA, 0.2% Triton-X 100, 1 mM DTT, and the bound protein was eluted with 0.5 mg ml<sup>-1</sup> 3 × Flag peptide. For anti-HA immunoprecipitation, total protein was pre-cleared with protein A agarose (Millipore) for 1 h, followed by precipitation with 2 μg anti-HA antibody (TianGen) together with protein A agarose for 4 h. Immunoprecipitates were separated by a 10% NuPAGE gel (Invitrogen) and detected by anti-HA or anti-Flag immunoblot.

**In vitro kinase assay.** His-*RIPK*, His-*BIK1* and GST-*AvrAC* were co-expressed in *E. coli* (strain BL21), and the His-tagged protein was purified by Ni-NTA affinity chromatography. Kinase assays were performed in 25 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 100 μM ATP, 1 mM DTT, γ-<sup>32</sup>P ATP (5 μCi), and 3 μg recombinant protein in a total volume of 20 μl. Reactions were allowed to proceed for 30 min at 30 °C and terminated by adding 5 × protein loading buffer and boiling for 5 min at 100 °C. The proteins were separated on a 12% SDS-PAGE gel, and incorporated radiolabel was visualized by autoradiography. As an equal loading control for proteins used in the kinase reactions, the SDS-PAGE gels were stained by Coomassie blue.

**BIK1 phosphorylation and migration shift assay.** *Arabidopsis* protoplasts were transfected with HA-tagged *BIK1* or *BIK1* mutants<sup>7</sup> together with *AvrAC*-Flag or truncation mutants of *AvrAC*, treated with 1 μM flg22 and total protein was extracted at various time points. Samples were separated by 10% SDS-PAGE gels followed by anti-HA immunoblot. For phosphatase treatment, total protein from protoplasts or affinity-purified recombinant protein was treated with λ protein phosphatase (New England Biolabs) according to the manufacturer's instructions.

**In vitro uridylylation assay.** GST-tagged AvrAC derivative proteins, His-tagged RIPK, BIK1 and PBL1 proteins were affinity purified and directly used for the uridylylation assay. The GST-PBS1 recombinant protein (a gift from J. Chai) was cleaved with PreScission Protease (GE) to remove GST before the enzymatic assay. Approximately 400 ng of GST-AvrAC, GST-ΔCTR, GST-CTR, GST-ΔFic or GST-AvrAC(H469A) were incubated with 2 μg of His-RIPK, His-BIK1, PBS1, His-PBL1, or His-tagged kinase domain of FLS2 or BAK1 (ref. 39) in 20 μl reaction buffer containing 25 mM Tris-HCl (pH 7.5), 25 mM MgCl<sub>2</sub>, 500 μM UTP, 1 mM DTT, α-<sup>32</sup>P UTP (5 μCi) for 30 min at 30 °C. Reactions were stopped by addition of loading buffer. The products were separated on 12% NuPAGE gels (Invitrogen) and visualized by autoradiography.

**Mass measurement of the recombinant AvrAC protein.** An affinity-purified GST-AvrAC protein solution was loaded into an analytical capillary column (75 μm × 2 cm) packed with Poros 20 R2 packing material (Applied Biosystems). An Agilent 1100 binary pump was used to generate HPLC gradient as follows: 0–100% B in 60 min (A = 0.1 M acetic acid in water; B = 0.1 M acetic acid/70% acetonitrile). The eluted proteins were sprayed into a QSTAR XL mass spectrometer (AB Sciex). The spray voltage was set at 2,100 V and the data was acquired in MS mode. The protein charge envelop from the raw spectrum was de-convoluted into non-charged form by the BioAnalyst software provided by the manufacturer.

**Mapping UMP modification sites.** Affinity-purified GST-AvrAC, His-BIK1, His-RIPK, BIK1-Flag and RIPK-Flag proteins were separated by SDS-PAGE. Protein bands on the SDS-PAGE gel were de-stained, and then reduced in 10 mM DTT at 56 °C for 30 min followed by alkylation in 55 mM iodoacetamide at dark for 1 h. After that the protein bands were in-gel digested with sequencing grade trypsin (10 ng μl<sup>-1</sup> trypsin, 50 mM ammonium bicarbonate, pH 8.0) overnight at 37 °C. Peptides were extracted with 5% formic acid/50% acetonitrile and 0.1% formic acid/75% acetonitrile sequentially and then concentrated to ~20 μl. The extracted peptides were separated by an analytical capillary column (50 μm × 10 cm) packed with 5 μm spherical C18 reversed phase material (YMC). An Agilent 1100 binary pump was used to generate the HPLC gradient as follows: 0–5% B in 5 min, 5–40% B in 25 min, 40–100% B in 15 min (A = 0.1 M acetic acid in water; B = 0.1 M acetic acid/70% acetonitrile). The eluted peptides were sprayed into a QSTAR XL mass spectrometer equipped with a nano-ESI ion source. The mass

spectrometer was operated in information-dependent mode with one MS scan followed by three MS/MS scans for each cycle. Database searches were performed on an in-house Mascot server (Matrix Science) and the following variable modifications were included: oxidation on methionine, carbamidomethylation on cysteine, UMP modification on serine or threonine residues, AMP modification on serine, threonine or tyrosine residues.

29. Wang, L. F., Tang, X. Y. & He, C. Z. The bifunctional effector AvrXccC of *Xanthomonas campestris* pv. *campestris* requires plasma membrane-anchoring for host recognition. *Mol. Plant Pathol.* **8**, 491–501 (2007).
30. Yuan, J. & He, S. Y. The *Pseudomonas syringae* Hrp regulation and secretion system controls the production and secretion of multiple extracellular proteins. *J. Bacteriol.* **178**, 6399–6402 (1996).
31. Li, X. *et al.* Flagellin induces innate immunity in nonhost interactions that is suppressed by *Pseudomonas syringae* effectors. *Proc. Natl Acad. Sci. USA* **102**, 12990–12995 (2005).
32. Schäfer, A. *et al.* Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69–73 (1994).
33. Zuo, J., Niu, Q. W. & Chua, N. H. An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.* **24**, 265–273 (2000).
34. Zhang, J. *et al.* A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host Microbe* **1**, 175–185 (2007).
35. Schwessinger, B. *et al.* Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1. *PLoS Genet.* **7**, e1002046 (2011).
36. Macho, A. P. *et al.* Competitive index in mixed infections: a sensitive and accurate assay for the genetic analysis of *Pseudomonas syringae*–plant interactions. *Mol. Plant Pathol.* **8**, 437–450 (2007).
37. Dow, J. M. *et al.* Biofilm dispersal in *Xanthomonas campestris* is controlled by cell-cell signaling and is required for full virulence to plants. *Proc. Natl Acad. Sci. USA* **100**, 10995–11000 (2003).
38. Schechter, L. M. *et al.* *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a Cya translocation reporter. *J. Bacteriol.* **186**, 543–555 (2004).
39. Cheng, W. *et al.* The AvrPtoB-BAK1 complex reveals two structurally similar kinase-interacting domains in a single type III effector. *Cell Host Microbe* **10**, 616–626 (2011).



# A novel putative auxin carrier family regulates intracellular auxin homeostasis in plants

Elke Barbez<sup>1,2</sup>, Martin Kubeš<sup>3</sup>, Jakub Rolčík<sup>4</sup>, Chloé Béziat<sup>1,2</sup>, Aleš Pěnčík<sup>5</sup>, Bangjun Wang<sup>6</sup>, Michel Ruiz Rosquete<sup>1,2</sup>, Jinsheng Zhu<sup>6</sup>, Petre I. Dobrev<sup>3</sup>, Yuree Lee<sup>7</sup>, Eva Zažímalová<sup>3</sup>, Jan Petrásek<sup>3</sup>, Markus Geisler<sup>6</sup>, Jiří Friml<sup>1</sup> & Jürgen Kleine-Vehn<sup>1,2</sup>

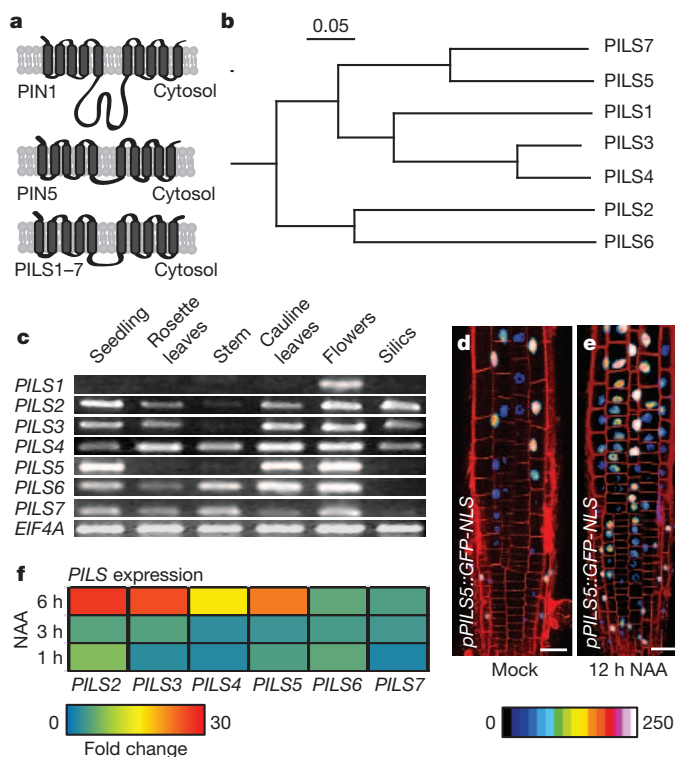
The phytohormone auxin acts as a prominent signal, providing, by its local accumulation or depletion in selected cells, a spatial and temporal reference for changes in the developmental program<sup>1–7</sup>. The distribution of auxin depends on both auxin metabolism (biosynthesis, conjugation and degradation)<sup>8–10</sup> and cellular auxin transport<sup>11–15</sup>. We identified *in silico* a novel putative auxin transport facilitator family, called PIN-LIKES (PILS). Here we illustrate that PILS proteins are required for auxin-dependent regulation of plant growth by determining the cellular sensitivity to auxin. PILS proteins regulate intracellular auxin accumulation at the endoplasmic reticulum and thus auxin availability for nuclear auxin signalling. PILS activity affects the level of endogenous auxin indole-3-acetic acid (IAA), presumably via intracellular accumulation and metabolism. Our findings reveal that the transport machinery to compartmentalize auxin within the cell is of an unexpected molecular complexity and demonstrate this compartmentalization to be functionally important for a number of developmental processes.

Prominent auxin carriers with fundamental importance during plant development are PIN-FORMED (PIN) proteins<sup>1–3,6,9,15</sup>. PIN1-type auxin carriers regulate the directional intercellular auxin transport at the plasma membrane. In contrast, atypical family member PIN5 regulates intracellular auxin compartmentalization into the lumen of the endoplasmic reticulum and its role in auxin homeostasis was recently identified<sup>15,16</sup>. PIN proteins have a predicted central hydrophilic loop, flanked at each side by five transmembrane domains. We screened *in silico* for novel PIN-like putative carrier proteins with a predicted topology similar to PIN proteins (Fig. 1a and Supplementary Fig. 2) and identified a protein family of seven members (Fig. 1b) in *Arabidopsis thaliana*, which we designated as the PILS proteins. In contrast to the similarities in the predicted protein topology, PIN and PILS proteins do not show pronounced protein sequence identity (10–18%), which limits the identification of PILS proteins by conventional, reciprocal basic local alignment search tool (BLAST) approaches. However, the distinct PIN and PILS protein families contain both the InterPro auxin carrier domain which is an *in-silico*-defined domain, aiming to predict auxin transport function (<http://www.ebi.ac.uk/panda/InterPro.html>). The PILS putative carrier family is conserved throughout the whole plant lineage, including unicellular algae (such as *Ostreococcus tauri* and *Chlamydomonas reinhardtii*) (Supplementary Fig. 3) where PIN proteins are absent<sup>16</sup>, indicating that PILS proteins are evolutionarily older.

PILS genes are broadly expressed in various tissues (Fig. 1c) and *PILS2–PILS7* were transcriptionally upregulated by auxin application in wild-type seedlings (Fig. 1d–f and Supplementary Fig. 4), indicating a role in auxin-dependent processes. To investigate the potential function of the putative PILS auxin flux facilitators in plant development, we overexpressed PILS proteins using the constitutive, viral 35S promoter. Ectopic expression of PILS genes, such as *PILS1* or *PILS3*,

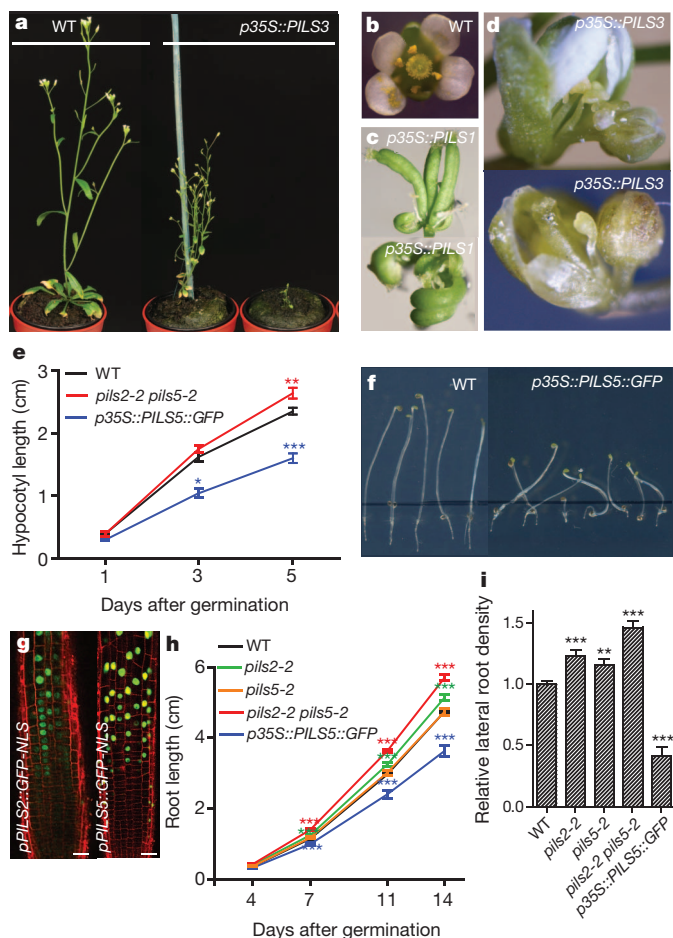
resulted in dwarfed and/or bushy plants showing severe defects in flower development, leading to sterility in the T1 generation (Fig. 2a–d). Flowers of these *PILS*-overexpressing plants showed severe patterning defects, such as homeotic transformation of flower organs into new flower buds, triplication of the gynoecium or unfused carpels (Fig. 2b–d). To circumvent sterility, we screened for weaker *p35S::PILS* lines and isolated moderately *PILS5* overexpressing lines showing fertile flower development.

To assess further the developmental importance and potential redundancy of PILS proteins in auxin-regulated processes, we focused



**Figure 1 | Novel PILS protein family.** **a**, Predicted topology of both PIN1-/PIN5-type PIN subfamilies and PILS proteins. **b**, Phylogenetic tree of the *A. thaliana* PILS proteins. Scale depicts 0.05 substitutions per position. **c**, Reverse transcription PCR (RT-PCR) of the seven PILS genes and *EIF4A* (control) in several plant tissues. **d**, **e**, 1-Naphthaleneacetic acid (NAA; 10 μM, 12 h)-induced *pPILS5::GFP-NLS* (GFP fused to nuclear localisation signal, NLS) expression in the root transition zone. Colour-code (black to white) depicts (low to high) GFP signal intensity. Propidium-iodide-stained cell walls in red. Scale bar, 25 μm. **f**, Quantitative RT-PCR of *PILS2–7* after 10 μM NAA treatment for 1, 3 and 6 h. Colour code from blue (low) to red (high) depicts fold changes (0–30) (see also Supplementary Fig. 4).

<sup>1</sup>Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Gent, Belgium. <sup>2</sup>Department of Applied Genetics and Cell Biology, University of Natural Resources and Life Sciences (BOKU), 1190 Vienna, Austria. <sup>3</sup>Institute of Experimental Botany, The Academy of Sciences of the Czech Republic, 16502 Praha 6, Czech Republic. <sup>4</sup>Laboratory of Growth Regulators, Faculty of Science, Palacký University and Institute of Experimental Botany AS CR, Šlechtitelů 11, 78371 Olomouc, Czech Republic. <sup>5</sup>Centre of the Region Haná for Biotechnological and Agricultural Research, Department of Growth Regulators, Faculty of Science, Palacký University, Šlechtitelů 11, 78371 Olomouc, Czech Republic. <sup>6</sup>Department of Biology - Plant Biology, University of Fribourg, Chemin du Musée 10, 1700 Fribourg, Switzerland. <sup>7</sup>Department of Plant Molecular Biology, University of Lausanne, Quartier Sorge, 1015 Lausanne, Switzerland.



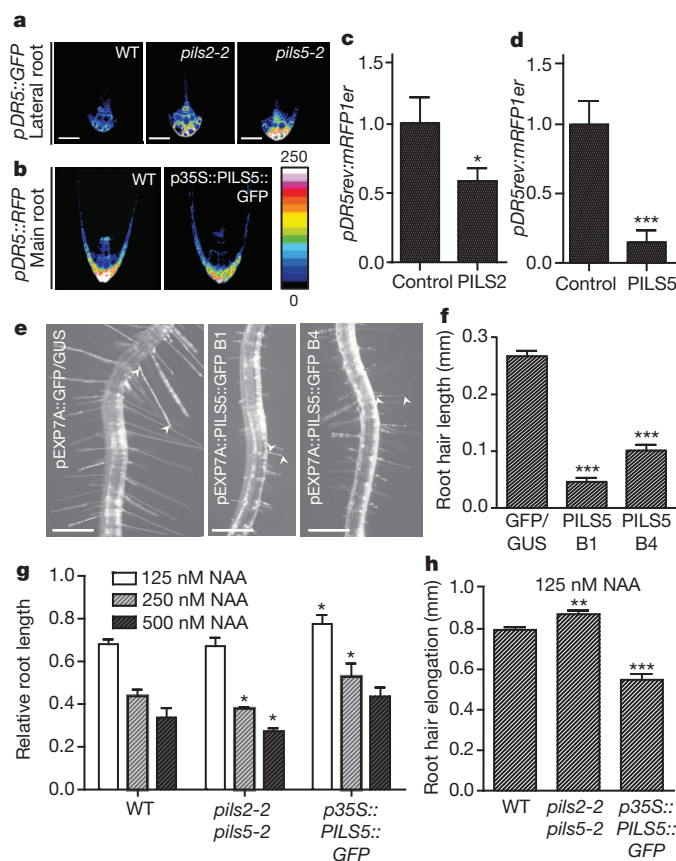
**Figure 2 | Phenotypes of PILS loss- and gain-of function mutants.** **a–d**, Wild-type (WT) and T1 transgenics strongly expressing *p35S::PILS*, such as *p35S::PILS3* (**a**), are sterile and show dwarf growth. **b–d**, Flowers of WT (**b**), *p35S::PILS1* (**c**) and *p35S::PILS3* (**d**). **e**, Hypocotyl length of dark-grown *pils2-2 pils5-2*, *p35S::PILS5::GFP* and WT seedlings ( $n > 15$ ). **f**, Dark-grown WT and *p35S::PILS5* seedlings (5 days after germination). **g**, *pPILS2::GFP-NLS* (left) and *pPILS5::GFP-NLS* (right) expression in the root transition zone (green). Propidium iodide-stained cell walls in red. Scale bar, 25  $\mu$ m. **h**, **i**, Root length (**h**) and lateral root density (14 days after germination) (**i**) of *pils2-2*, *pils5-2*, *pils2-2 pils5-2*, *p35S::PILS5::GFP* and WT seedlings ( $n > 40$ ). Error bars represent s.e.m. Student *t*-test *P*-values: \* $P < 0.05$ , \*\* $P < 0.001$ , \*\*\* $P < 0.0001$ .

on *PILS2* and *PILS5*, because they are the most abundantly expressed *PILS* genes in seedlings (Fig. 1c) and display partially overlapping expression domains (Fig. 2g). Initially, we investigated *PILS* function in auxin-dependent hypocotyl and root growth. Dark-grown *p35S::PILS5::GFP*-expressing (coding for a *PILS5*–green fluorescent protein (GFP) fusion) and *pils2 pils5* double-mutant seedlings showed reduced and enhanced hypocotyl growth, respectively (Fig. 2e and Supplementary Fig. 5). *PILS5* gain-of-function also resulted in agravitropic hypocotyl growth (Fig. 2f). *PILS2* and *PILS5* showed a particular overlapping expression in the root transition zone (Fig. 2g), suggesting a redundant role in regulating root growth. Indeed, *pils2* single-mutant and more pronounced *pils2 pils5* double-mutant seedlings showed significantly longer roots compared to wild-type seedlings, whereas seedlings overexpressing *PILS5* had shorter roots (Fig. 2h). Collectively, our data indicate the requirement of defined *PILS* protein activity for auxin-dependent growth regulation.

Beside the root and shoot organ growth, auxin tightly controls *de novo* organ formation such as lateral root organogenesis<sup>1</sup>. Intriguingly, *pils2* and *pils5* single-mutants and more pronounced *pils2 pils5* double-mutants showed higher lateral root density (Fig. 2i). On the contrary, *PILS5* gain-of-function reduced lateral rooting (Fig. 2i).

These findings indicate developmental importance of *PILS* proteins in auxin-regulated processes, such as *de novo* organ formation and growth regulation.

Next we investigated whether auxin responses are affected in *pils* mutants using the auxin response reporter DR5 (ref. 17). *pils2-2* knockdown and *pils5-2* knockout mutants did not show altered DR5 activity in the main root tips (Supplementary Fig. 6a), but showed higher *pDR5rev::GFP* signal intensity in lateral roots (Fig. 3a). Moderately *p35S::PILS5::GFP*-expressing seedlings showed a visibly reduced auxin response maximum in the very root tip of main (Fig. 3b) and lateral roots (Supplementary Fig. 6b). To distinguish between direct and indirect effects, we investigated the effect of *PILS* proteins on nuclear auxin signalling at the cellular level. Therefore, we transiently co-expressed *PILS2* or *PILS5* together with the auxin response reporter *pDR5rev::mRFP1er* (expressing the red fluorescent protein (RFP) in response to auxin signalling) in tobacco Bright Yellow 2 (BY-2) cells. *PILS2* or *PILS5* expression in BY-2 cells reduced auxin signalling as visualized by *pDR5rev::mRFP1er* activity (Fig. 3c, d). These findings indicate that *PILS* proteins negatively affect nuclear auxin signalling, presumably by affecting cellular auxin homeostasis.



**Figure 3 | PILS proteins affect auxin-dependent cellular growth.** **a**, **b**, DR5 promoter activity in *pils2-2* and *pils5-2* (lateral roots) (**a**) and *p35S::PILS5::GFP* (main root) (**b**) compared to wild type (WT) (see also Supplementary Fig. 6). Colour-code (black to white) depicts (low to high) *pDR5rev::GFP* and *pDR5rev::mRFP1er* signal intensity. Scale bars, 25  $\mu$ m. **c**, **d**, Mean grey value of *pDR5rev::mRFP1er* signal intensity in BY-2 cells expressing *p35S::PILS2* (**c**) or *p35S::PILS5* (**d**) compared to control cells (*p35S::GFP::HDEL*) ( $n > 60$  cells). **e**, **f**, Root hair length (**e**, arrowheads) and quantification (**f**) of transgenic lines expressing *pEXP7A::GFP::GUS* (control) and *pEXP7A::PILS5::GFP* (independent lines B1 and B4) ( $n = 20$  seedlings with 400 counted root hairs in total). Scale bars, 250  $\mu$ m. **g**, Relative root length of NAA-treated *pils2-2 pils5-2* and *p35S::PILS5::GFP* seedlings compared to WT ( $n > 20$ ). **h**, Auxin-induced root hair elongation ( $n = 20$  seedlings with 400 counted root hairs in total). Mean untreated root hair length was subtracted from the treated values. Error bars represent s.e.m. Student *t*-test *P*-values: \* $P < 0.05$ , \*\* $P < 0.001$ , \*\*\* $P < 0.0001$ .



To address whether cellular PILS action on auxin signalling affects cellular growth, we expressed *PILS1*, *PILS3* and *PILS5* under a root-hair-specific promoter. Deviations in free (active) indole-3-acetic acid (IAA) levels or in auxin signalling induce (high levels) or repress (low levels) root hair growth<sup>18</sup>. As expected, the root-hair-specific expression of *PILS1*, *PILS3* and *PILS5* significantly reduced root hair length (Fig. 3e, f and Supplementary Fig. 7a–c), possibly due to PILS-dependent regulation of auxin homeostasis and signalling.

Next we tested whether PILS proteins affect auxin-dependent cellular growth responses. We treated PILS gain- and loss-of-function mutants with auxin that inhibits primary root growth. *pils2 pils5* loss- and *PILS5* gain-of-function mutants showed hyper- and hyposensitive root growth, respectively (Fig. 3g). In agreement with these observations, the auxin-promoted root hair growth was also enhanced in *pils2 pils5* double-mutants and reduced in *PILS5* overexpressors (Fig. 3h). This set of data indicates that PILS putative auxin facilitators modulate auxin-dependent growth responses during plant development.

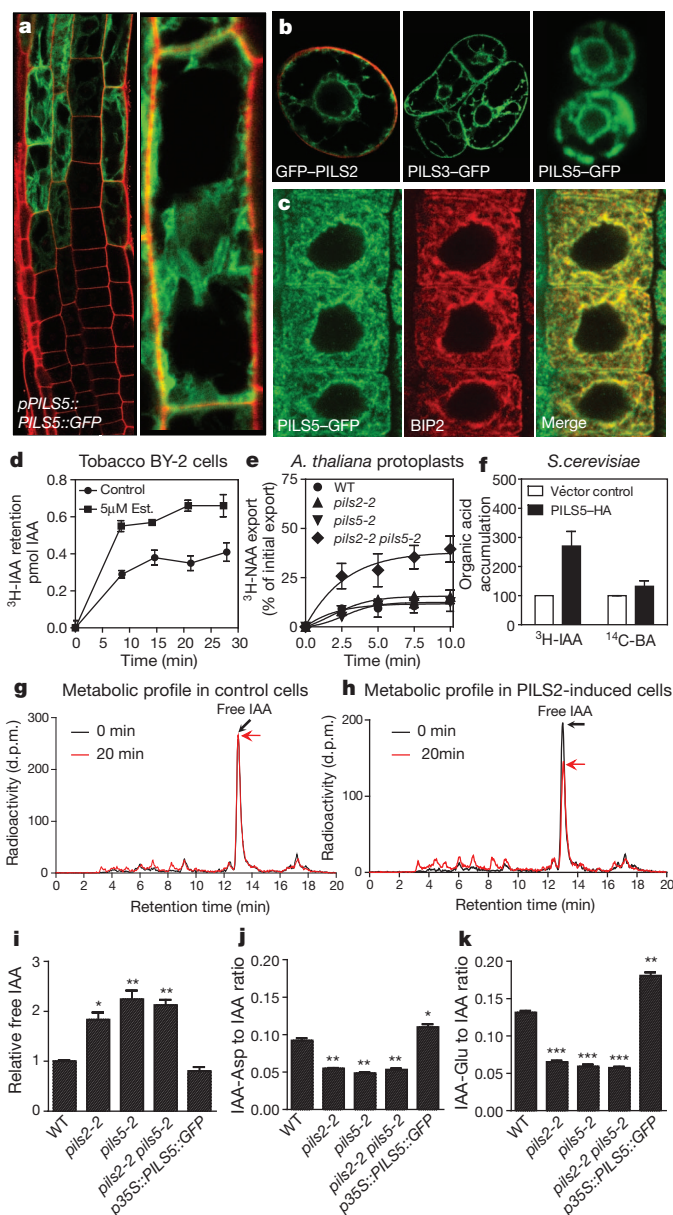
To unravel the mechanism by which PILS proteins regulate auxin-dependent plant development, we investigated the subcellular localization of PILS proteins. We introduced amino- or carboxy-terminal GREEN/RED FLUORESCENT PROTEIN (GFP/RFP) fusions with PILS proteins and transiently or stably expressed these fusion proteins in tobacco BY-2 cell culture, *Arabidopsis* seedlings and heterologously in yeast (Fig. 4a–c and Supplementary Fig. 8). *PILS1*, *PILS2*, *PILS3*, *PILS5*, *PILS6* and *PILS7* GFP/RFP fusion proteins localized to the endoplasmic reticulum in all analysed systems (Fig. 4a–c and Supplementary Figs 8 and 10c). In contrast, N- and C-terminal *PILS4* fusions did not show detectable fluorescence. The transgenic *pPILS5::PILS5::GFP* also showed endoplasmic reticulum localization (Fig. 4a) and complemented the *pils2 pils5* double-mutant to the *pils2* single-mutant level (Supplementary Fig. 9), indicating PILS function at the endoplasmic reticulum.

Next, we addressed whether putative PILS auxin carriers affect cellular auxin accumulation. We generated *PILS2* oestradiol-inducible tobacco BY-2 cell cultures and performed <sup>3</sup>H-IAA accumulation assays. *PILS2* induction increased the accumulation of radioactivity in BY-2 cells (Fig. 4d). In accordance with the auxin accumulation assays in BY-2 cells, *pils2 pils5* double-mutant protoplasts showed significantly higher auxin export (Fig. 4e and Supplementary Fig. 10a), indicating reduced auxin retention capacity in *pils2 pils5* loss-of-function mutants. These gain and loss-of-function studies consistently illustrate that *PILS2* and *PILS5* function at the endoplasmic reticulum controls cellular accumulation of auxin.

To additionally address PILS protein function in a non-plant system, we expressed *PILS3*, *PILS5* and *PILS7* (tagged to GFP or haemagglutinin (HA)) in *Saccharomyces cerevisiae* yeast cells. In accordance with the effect of *PILS2* and *PILS5* in plant cells, the expression of *A. thaliana PILS3*, *PILS5* and *PILS7* in yeast also increased retention of exogenously applied auxin (Fig. 4f and Supplementary Fig. 10b). To assess the specificity to auxin we used the common organic control benzoic acid (BA). Notably, *PILS3*, *PILS5* and *PILS7* did not affect the cellular accumulation of BA (Fig. 4f and Supplementary Fig. 10b). These findings indicate that PILS proteins specifically regulate cellular auxin accumulation.

PILS proteins increase cellular auxin accumulation, but decrease auxin signalling, which is reminiscent of the regulation of auxin metabolism shown for the PIN5 auxin transporter<sup>16</sup>. To assess the potential action of PILS proteins on auxin metabolism, we analysed the effect of PILS proteins on the auxin metabolism using the inducible *PILS2* BY-2 cell line. The high-performance liquid chromatography (HPLC) chromatogram of IAA showed a *PILS2*-dependent reduction of free IAA within 20 min (Fig. 4g, h), indicating that *PILS2* activity affects auxin metabolism.

In accordance with the auxin metabolite profiling in BY-2 cell cultures, *pils2*, *pils5* and *pils2 pils5* mutant *Arabidopsis* seedlings showed significantly higher free IAA levels compared to wild-type seedlings (Fig. 4i). Intriguingly, the ratio of amid auxin conjugates IAA-Glutamate (Glu)



**Figure 4 | PILS involvement in cellular auxin homeostasis.** **a–c**, GFP–*PILS2* (**a**), *PILS3*–GFP (**b**), *PILS5*–GFP (**a–c**) under endogenous (**a**) or constitutive promoter (**b**, **c**) in *A. thaliana* root (**a**, **c**), tobacco BY-2 cells (**b**, left panels) and *Saccharomyces cerevisiae* yeast cells (**b**, right panel). **c**, Immunocytochemistry of *PILS5*–GFP and the endoplasmic reticulum marker BIP2. **d**, <sup>3</sup>H-IAA retention in tobacco BY-2 cells upon oestradiol (Est.)-induced *PILS2* expression ( $n = 3$  repetitions). **e**, <sup>3</sup>H-NAA export assay in *A. thaliana* protoplasts of wild type (WT), *pils2-2*, *pils5-2* and *pils2 pils5* mutants ( $n = 3$  repetitions; see also Supplementary Fig. 10a). **f**, Accumulation of auxin and benzoic acid in *Saccharomyces cerevisiae* yeast cells transformed with *pGPD::PILS5::HA* or an empty vector ( $n = 3$  repetitions; see also Supplementary Fig. 10b). **g**, **h**, <sup>3</sup>H-IAA metabolic profile (HPLC chromatogram) in *PILS2* non-oestradiol-induced (**g**) and oestradiol-induced (**h**) tobacco BY-2 cells after 0 (black line) and 20 min (red line). **i–k**, Liquid chromatography-mass spectrometry (LC-MS)-derived free IAA levels (**i**) and ratios of IAA-Asp (**j**) and IAA-Glu (**k**) to free IAA in WT, *pils2-2*, *pils5-2*, *pils2-2 pils5-2* and *p35S::PILS5::GFP* ( $n = 3$  repetitions). Error bars represent s.e.m. Student *t*-test *P*-values: \* $P < 0.05$ , \*\* $P < 0.001$ , \*\*\* $P < 0.0001$ . Scale bar, 10 μm.

and IAA-Aspartate (Asp) to free IAA was also significantly shifted towards free IAA in *pils2* and *pils5* loss-of-function mutants (Fig. 4j, k). Moderately *p35S::PILS5::GFP*-expressing seedlings showed a visibly reduced auxin response maximum in the very root tip (Fig. 3b), but only a mild and statistically non significant reduction in free IAA levels



at the whole seedling level (Fig. 4i). However, the ratio of IAA-Glu and IAA-Asp to free IAA was significantly shifted towards the conjugates in *p35S::PILS5::GFP* seedlings (Fig. 4j, k), indicating a higher rate of auxin conjugation.

These findings are indicative of a PILS function in cellular auxin homeostasis by regulating auxin metabolism. Hence, we propose a model in which PILS proteins at the endoplasmic reticulum membrane facilitate intracellular auxin accumulation, which seems to contribute to the possibly compartmentalized regulation of auxin metabolism (Supplementary Fig. 1). It is tempting to speculate that auxin conjugation could also take place in the endoplasmic reticulum, although the molecular components remain to be identified.

Our *in silico* and reverse genetics approaches led to the identification of a novel family of putative auxin transport facilitators. All our genetic, pharmacological, cell biological, physiological and biochemical approaches consistently suggest that PILS proteins function at the endoplasmic reticulum membrane, regulate intracellular auxin accumulation and affect free IAA levels, presumably through conjugation-based auxin metabolism. The PILS action on cellular auxin homeostasis is reminiscent of the function of the atypical PIN family member PIN5 (ref. 16). We uncovered that an additional, distinct protein family regulates intracellular auxin homeostasis. Our analyses of the PILS proteins suggest that intracellular auxin transport and, hence, auxin compartmentalization might be evolutionarily older than directional, cell-to-cell PIN-dependent auxin transport mechanisms. The identification of a novel protein family for the regulation of intracellular auxin homeostasis highlights the evolutionary and developmental importance of intracellular auxin transport. Further studies will address the potential interplay or possible diversified function of endoplasmic reticulum localized PIN5 and PILS1–7 proteins.

## METHODS SUMMARY

**Plant material and growth conditions.** We used *Arabidopsis thaliana* of ecotype Columbia 0 (Col-0). The *Nicotiana tabacum* L. cv. Bright Yellow 2 (BY-2) cell line<sup>19</sup> was used as suspension-cultured cells. *pils2-1* (SALK\_024808), *pils2-2* (SALK\_125391), *pils5-1* (SALK\_070653) and *pils5-2* (SALK\_072996) were obtained from the Nottingham *Arabidopsis* Stock Centre (NASC). Gateway cloning was used to construct *pPILS2::GFP/GUS*, *pPILS5::GFP/GUS*, *p35S::PILS1-7*, *p35S::GFP::PILS1-7*, *p35S::PILS1-7::GFP*, *pPILS5::PILS5::GFP* and *pMDC7\_B(pUBQ)::PILS2*. The following lines and constructs have been previously described: *pDR5rev::mRFP1er<sup>20</sup>*, *pDR5rev::GFP<sup>2</sup>* and *p35S::GFP::HDEL<sup>21</sup>*. Seeds were stratified at 4 °C for 2 days in the dark. Seedlings were grown vertically on half Murashige and Skoog medium. Plants were grown under long-day (16 h light/8 h dark) conditions at 20–22 °C.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 8 December 2011; accepted 24 February 2012.

Published online 15 April 2012.

- Benková, E. *et al.* Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* **115**, 591–602 (2003).
- Friml, J. *et al.* Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* **426**, 147–153 (2003).
- Reinhardt, D. *et al.* Regulation of phyllotaxis by polar auxin transport. *Nature* **426**, 255–260 (2003).

- Leyser, O. Dynamic integration of auxin transport and signalling. *Curr. Biol.* **16**, R424–R433 (2006).
- Dubrovsky, J. G. *et al.* Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. *Proc. Natl Acad. Sci. USA* **105**, 8790–8794 (2008).
- Sorefan, K. *et al.* A regulated auxin minimum is required for seed dispersal in *Arabidopsis*. *Nature* **459**, 583–586 (2009).
- Prasad, K. *et al.* *Arabidopsis* PLETHORA transcription factors control phyllotaxis. *Curr. Biol.* **21**, 1123–1128 (2011).
- Woodward, A. W. & Bartel, B. Auxin: regulation, action, and interaction. *Ann. Bot.* **95**, 707–735 (2005).
- Ikeda, Y. *et al.* Local auxin biosynthesis modulates gradient-directed planar polarity in *Arabidopsis*. *Nature Cell Biol.* **11**, 731–738 (2009).
- Zhao, Y. Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* **61**, 49–64 (2010).
- Bennett, M. J. *et al.* *Arabidopsis* AUX1 gene: a permease-like regulator of root gravitropism. *Science* **273**, 948–950 (1996).
- Luschnig, C., Gaxiola, R. A., Grisafi, P. & Fink, G. R. EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in *Arabidopsis thaliana*. *Genes Dev.* **12**, 2175–2187 (1998).
- Geisler, M. *et al.* Cellular efflux of auxin catalyzed by the *Arabidopsis* MDR/PGP transporter AtPGP1. *Plant J.* **44**, 179–194 (2005).
- Petrášek, J. *et al.* PIN proteins perform a rate-limiting function in cellular auxin efflux. *Science* **312**, 914–918 (2006).
- Zažímalová, E., Murphy, A. S., Yang, H., Hoyerová, K. & Hošek, P. Auxin transporters—why so many? *Cold Spring Harb. Perspect. Biol.* **2**, a001552 (2010).
- Mravec, J. *et al.* ER-localized PIN5 auxin transporter mediates subcellular homeostasis of phytohormone auxin. *Nature* **459**, 1136–1140 (2009).
- Ulmasov, T., Murfett, J., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**, 1963–1971 (1997).
- Lee, S. H. & Cho, H. T. PINOID positively regulates auxin efflux in *Arabidopsis* root hair cells and tobacco cells. *Plant Cell* **18**, 1604–1616 (2006).
- Nagata, T., Nemoto, Y. & Hasegawa, S. Tobacco BY-2 cell line as the “HeLa” cells in the cell biology of higher plants. *Int. Rev. Cytol.* **132**, 1–30 (1992).
- Marin, E. *et al.* miR390, *Arabidopsis* TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell* **22**, 1104–1117 (2010).
- Langhans, M. *et al.* *In vivo* trafficking and localization of p24 proteins in plant cells. *Traffic* **9**, 770–785 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to C. Braeckman for plant transformation; W. Ardiles for sequencing support; L. Charrier for technical assistance; A. Maizel, N. Geldner and P. Pimpl for providing material; J.K.-V. group members for critical reading of the manuscript and the BOKU-VIBT Imaging Center for access and expertise. This work was supported by the Vienna Science and Technology Fund (WWTF) (to J.K.-V.), the Agency for Innovation by Science and Technology (IWT) (predoctoral fellowship to E.B.), the Odysseus program of the Research Foundation Flanders (to J.F.), the Swiss National Funds (to M.G.), the Ministry of Education, Youth and Sports of the Czech Republic (LC06034) (to E.Z.), Grant Agency of the Czech Republic project P305/11/2476 (to J.P.) and P305/11/0797 (to E.Z.).

**Author Contributions** E.B. and J.K.V. conceived the project. E.B. carried out most of the experiments. M.K., P.I.B., E.Z. and J.P. performed auxin metabolite profile and auxin accumulation in BY-2. C.B. analysed auxin-dependent PILS expression and contributed to phenotype analysis. M.R.R. contributed to PILS cloning. J.R. and A.P. measured auxin content in *Arabidopsis*. B.W., J.Z. and M.G. performed auxin accumulation in yeast and protoplasts. Y.L. modified the oestradiol-inducible vector. E.B., M.K., J.R., E.Z., J.P., M.G., J.F. and J.K.V. discussed the experimental procedures. All authors analysed and discussed the data; E.B. and J.K.V. wrote the paper and all authors saw and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.K.-V. ([juergen.kleine-vehn@boku.ac.at](mailto:juergen.kleine-vehn@boku.ac.at)).

## METHODS

**PILS gene accession codes.** Sequence data from this article can be found in The Arabidopsis Information Resource (TAIR; <http://www.arabidopsis.org/>) or GenBank/EMBL databases under the following accession numbers: *PILS1* (At1g20925), *PILS2* (At1g71090), *PILS3* (At1g76520), *PILS4* (At1g76530), *PILS5* (At2g17500), *PILS6* (At5g01990), *PILS7* (At5g65980).

**Plant material, growth conditions and DNA constructs.** We used *Arabidopsis thaliana* of ecotype Columbia 0 (Col-0). The *Nicotiana tabacum* L. cv. Bright Yellow 2 (BY-2) cell line<sup>19</sup> was used as suspension-cultured cells. *pils2-1* (SALK\_024808), *pils2-2* (SALK\_125391), *pils5-1* (SALK\_070653) and *pils5-2* (SALK\_072996) were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Insertion sites were verified, homozygous lines selected and the decrease or absence of the respective *PILS* transcript was shown by RT-PCR. The *pils2-2* and the *pils5-2* mutants were crossed into *DR5rev::GFP*. Gateway cloning was used to construct *pPILS2::GFP::GUS*, *pPILS5::GFP::GUS*, *p35S::PILS1-7*, *p35S::GFP::PILS1-7*, *p35S::PILS1-7::GFP*, *pPILS5::PILS5::GFP* and *pMDC7\_B(pUBQ)::PILS2*. The *PILS* full-length genomic fragments, complementary DNA and promoter regions were amplified by PCR from genomic DNA and cDNA, respectively. The PCR was performed using the high fidelity DNA polymerase "I proof" (Bio-Rad). The primers used are given below. The full genomic and cDNA fragments were cloned into the pDONR221 (Invitrogen) vector and the promoter regions into pDONR-P4P1 using Invitrogen BP-clonase according to manufacturer's instructions. Coding sequences were transferred from the entry clones to gateway-compatible destination vectors (given below) using the Invitrogen LR clonase(+) according to manufacturer's instructions. The resulting constructs were transformed into Col-0 plants by floral dipping in *Agrobacterium tumefaciens* liquid cultures. Yeast vectors were transformed into budding yeast (*Saccharomyces cerevisiae*) via electroporation. The *p35S::PILS5::GFP* line was crossed into *pDR5rev::mRFP1er*<sup>20</sup>. The following lines and constructs have been previously described: *pDR5rev::mRFP1er*<sup>20</sup>, *pDR5rev::GFP*<sup>2</sup> and *p35S::GFP::HDEL*<sup>21</sup>. Seeds were stratified at 4 °C for 2 days in the dark. Seedlings were grown vertically on half Murashige and Skoog medium. Plants were grown under long-day (16 h light/8 h dark) conditions at 20–22 °C.

**Chemicals.** 1-Naphthaleneacetic acid (NAA) was supplied by Duchefa, 2,4-dichlorophenoxy acetic acid, oestradiol and propidium iodide by Sigma-Aldrich and <sup>3</sup>H-indole-3-acetic acid (<sup>3</sup>H-IAA), <sup>3</sup>H-naphthalene-1-acetic acid (<sup>3</sup>H-NAA) and <sup>14</sup>C-benzoic acid (<sup>14</sup>C-BA) (specific radioactivity 20 Ci mmol<sup>-1</sup>) by American Radiolabelled Chemicals.

**RNA extraction and quantitative real time PCR (qPCR).** Whole RNA of seedlings was extracted using the RNeasy Mini Kit (Qiagen) in technical triplicates, the extracted RNA samples were treated with DNase (Ambion). qPCR analysis was performed using ICycler (Bio-Rad) with the Platinum SYBR Green qPCR Super-UDG kit (Invitrogen) following recommendations of the manufacturer. qPCR was carried out in 96-well optical reaction plates heated for 10 min to 95 °C to activate hot-start *Taq* DNA polymerase, followed by 40 cycles of denaturation for 60 s at 95 °C and annealing-extension for 60 s at 58 °C. Target quantifications were performed with specific primer pairs (given later) designed using Beacon Designer 4.0 (Premier Biosoft International). Expression levels were normalized to the expression levels of translation initiation factor EIF4A. The primers used are given later.

**Phenotypic analysis.** For analysis of the root length and lateral root density, plates were scanned on a flat-bed scanner. Root hairs were imaged with a binocular microscope (Leica). For hypocotyls analysis, seeds on plates were exposed to light for 3 h at 18 °C, and cultivated in the dark at 20 °C. Seedlings were imaged in real time with an infrared camera (Canon) to define the exact moment of germination and analysed 1, 3 and 5 days after germination. Hypocotyls, root and root hair lengths were measured with the ImageJ (<http://rsb.info.nih.gov/ij/>) software. Lateral root density for each seedling was obtained by calculating the number of lateral roots per root length unit 14 days after germination. For analysis of hypocotyls length, a minimum of 15 hypocotyls per condition or mutant line were analysed in each experiment. For analysis of root length and lateral root density, a minimum of 40 plants per condition or mutant line were analysed in each experiment. Means and standard errors were calculated and the statistical significance was evaluated by the student *t*-test using the GraphPad Prism5 (<http://www.graphpad.com>) software. For the analysis of root hair growth, 20 seedlings per transgenic line were imaged by binocular (Leica) and 20 root hairs (randomly chosen in the root hair zone) per seedling were measured with the ImageJ (<http://rsb.info.nih.gov/ij/>) software. The mean and standard error of the mean per transgenic line were calculated and the statistical significance was evaluated by the student *t*-test. To obtain the auxin-dependent root hair elongation, the same number of root hairs per seedling, seedlings per mutant line and condition were analysed as described above. The untreated mean average root hair length of the respective genotype was subtracted from the individual auxin-treated root hair length to obtain auxin induced growth in millimetre. The mean and standard error of the mean of the respective genotype were calculated and the

statistical significance was evaluated by the student *t*-test using the GraphPad Prism5 (<http://www.graphpad.com>) software. All experiments were performed in at least three independent biological repetitions.

**BY-2 plant material.** Cells of tobacco line BY-2 (*Nicotiana tabacum* L., cv. Bright Yellow 2)<sup>19</sup> transformed with *pMDC7\_B(pUBQ)::PILS2* were cultured in liquid cultivation medium (3% (w/v) sucrose, 4.3 g l<sup>-1</sup> Murashige and Skoog salts, 100 mg l<sup>-1</sup> inositol, 1 mg l<sup>-1</sup> thiamine, 0.2 mg l<sup>-1</sup> 2,4-dichlorophenoxy acetic acid, and 200 mg l<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub> (pH 5.8)). BY-2 cell lines were cultivated in darkness at 26 °C on an orbital incubator (Sanyo Gallenkamp, Schöeller Instruments; 150 r.p.m., 32-mm orbit) and subcultured weekly. Stock BY-2 calli were maintained on media solidified with 0.6% (w/v) agar and subcultured monthly.

**Transient transformation of BY-2 cells and monitoring of cellular auxin signalling in BY2.** Ten ml of three-day-old cells were harvested on filter paper by vacuum filtration and kept on plates with solid BY-2 medium. The cells were transformed via particle bombardment with a PDS 1000/He biolistic system (Bio-Rad) according to the manufacturer's instructions ([http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin\\_9075.pdf](http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_9075.pdf)). 2 µl of plasmid DNA (0.05 µg µl of the *pDR5rev::mRFP1er* construct and 0.1 µg µl<sup>-1</sup> of *p35S::PILS2* and *p35S::PILS5*) was added to 6.25 µl of 1.6-µm diameter gold particles (dissolved in 50% glycerol). The suspension was supplemented with 2.5 µl spermidine (0.1 M stock solution) and 6.25 µl CaCl<sub>2</sub> (2.5 M stock solution). The particles were pelleted by centrifugation, washed twice with 70% and 100% ethanol and, subsequently, resuspended in 10 µl of 100% ethanol. Cells were bombarded under a pressure of 1,100 pounds per square inch. The plates were sealed with Parafilm and kept in the dark for 18 h at 25 °C. For microscopic analysis, cells were gently transferred from the filter to a microscopy slide (in water) and subsequently covered with a cover slip. Samples were analysed via confocal microscopy. The *pDR5rev::mRFP1er* expression was evaluated by defining the mean grey value (MGV) of each imaged cell (middle sections). For each experiment, confocal settings were defined based on the *pDR5rev::mRFP1er* signal of the control cells and remained unchanged during the respective experiments. Transformants were identified on the basis of the fluorescence of both proteins and imaged with a ×40 objective. Every experiment/transformation was done in triplicate and for each condition a total number of at least 60 transformed cells were imaged. For each experiment, the means and standard errors were calculated and the statistical significance (independence between the two populations) was obtained by the student *t*-test using the GraphPad Prism5 (<http://www.graphpad.com>) software.

**Immunocytochemistry.** Whole-mount immunological staining on 5-day-old seedlings was done in an Intavis robot according to the described protocol<sup>22</sup>. The antibodies used at the final dilutions were monoclonal mouse anti-BIP2 (Hsc70) at 1:200 (Stressgen Bioreagents), monoclonal rabbit anti-GFP at 1:600 (Invitrogen). The secondary anti-mouse (Invitrogen) and anti-rabbit (Sigma-Aldrich) antibodies conjugated with Cy3 and Alexa488 respectively were used at 1:600 dilution.

**Microscopy.** Confocal microscopy was done with a Zeiss 710 microscope (Zeiss) or Leica SP5 (Leica). Fluorescence signals for GFP (excitation 488 nm, emission peak 509 nm), mRFP1 (excitation 561 nm, emission peak 607 nm) and propidium iodide (PI) staining (excitation 536 nm, emission peak 617 nm) were detected with a ×20, ×40 (water immersion) or ×63 (water immersion) objective. Sequential scanning was used for double labelling to avoid crosstalk between channels. Fluorescence signal intensity was analysed with ImageJ (<http://rsb.info.nih.gov/ij/>) software and data were statistically evaluated with Microsoft Excel 2007.

**Auxin transport assays in tobacco BY-2 cells, baker's yeast and Arabidopsis thaliana protoplasts.** Auxin accumulation with 2-day-old BY-2 cells was measured as previously described<sup>14,23</sup>. The <sup>3</sup>H-IAA was added to give a final concentration of 2 nM. Accumulation results were expressed as pmols of particular auxin accumulated per 10<sup>6</sup> cells. The 0.5-ml aliquots of cell suspension were collected continuously and accumulation of label was terminated by rapid filtration under reduced pressure on 22-mm-diameter cellulose filters. The cell cakes and filters were transferred to scintillation vials, extracted in 0.5 ml of 96% ethanol for 30 min, and afterwards 4 ml of scintillation solution (EcoLite Liquid Scintillation Fluid, MP Biomedicals) was added. Radioactivity was determined by liquid scintillation counter Packard Tri-Carb 2900TR (Packard-Canberra, Meridian). Yeast <sup>3</sup>H-IAA loading was quantified with the unspecific <sup>14</sup>C-benzoic acid as control assayed in parallel and performed as previously described<sup>24</sup>. Relative export is calculated from yeast-retained radioactivity as follows: ((radioactivity in the yeast at time *t* = 10 min) – (radioactivity in the yeast at time *t* = 0)) × (100%)/(radioactivity in the yeast at *t* = 0 min). Unspecific loading due to diffusion was eliminated by vector control subtraction. IAA export from *Arabidopsis thaliana* mesophyll protoplasts was analysed as described<sup>16</sup>.

**HPLC metabolic profiling in tobacco BY-2 cells.** Two-days-old BY-2 cells were prepared for the experiment by equilibration in uptake buffer as already described for accumulation assays<sup>14</sup>. Experiments were done in uptake buffer and under

standard cultivation conditions. Cells were incubated with addition of 20 nM <sup>3</sup>H-IAA for a period of 0 and 20 min. Cells and media (uptake buffer) were collected and frozen in liquid nitrogen (100 mg of fresh weight and 5 ml per sample). Extraction and purification of auxin metabolites in cells and media were performed as described<sup>25,26</sup>. The metabolites were separated on HPLC consisting of autosampler and 235C diode array detector (Perkin Elmer), column Luna C18 (2), 150 × 4.6 mm, 3 μm (Phenomenex, Torrance, USA), mobile phase A: 40 mM CH<sub>3</sub>COONH<sub>4</sub>, (pH 4.0) and mobile phase B: CH<sub>3</sub>CN/CH<sub>3</sub>OH, 1/1, (v/v). Flow rate was 0.6 ml min<sup>-1</sup> with linear gradient 30–50% B for 10 min, 50–100% B for 1 min, 100% B for 2 min, 10–30% B for 1 min. The column eluate was monitored on 235C DAD followed by Ramona 2000 flow-through radioactivity detector (Raytest GmbH) after online mixing with three volumes (1.8 ml min<sup>-1</sup>) of liquid scintillation cocktail (Flo-Scint III, Perkin Elmer). The radioactive metabolites were identified on the basis of comparison of their retention times with authentic standards. For the results presentation the total integrated area of chromatogram plots has been normalized based on the equalization of total accumulated radiolabel.

**In silico and phylogenetic analysis.** PILS genes were identified via the SMART-protein tool from EMBL (<http://smart.embl-heidelberg.de/>)<sup>27,28</sup>. Phylogenetic tree of AtPILS was constructed with the DNA-man software version 4.0. PILS topologies were defined by the online HMM-top tool (<http://www.enzim.hu/hmmtop/>)<sup>29</sup> and visualized by the TMRPres2D software (<http://biophysics.biol.uoa.gr/TMRPres2D/download.jsp>)<sup>30</sup>. PILS orthologues were identified with the online tool Plaza (<http://bioinformatics.psb.ugent.be/plaza/>)<sup>31</sup>.

**Free IAA and conjugate measurements in *Arabidopsis thaliana*.** For the quantification of free IAA and its amino acid conjugates, approximately 10 mg of plant material was taken into analysis. The samples were processed as previously described<sup>32</sup> and quantified by UHPLC-MS/MS.

**Used primers and vectors.** Genotyping primers: pils2-1 RP, CTGGAGAAACC TGACATCTCG; pils2-1 LP, GATTGAAGCCGGCTTAAATTC; pils2-2 RP, CTGGAGAAACCTGACATCTCG; pils2-2 LP, TACCATTTGATCTGTCTTCG GG; pils5-1 RP, TTGAGACCCGTATCATTGGAG; pils5-1 LP, TGTCTTG ATAAACCTTTTCAGG; pils5-2 RP, TACTGCACCGAAAATGAAACC; pils5-2 LP, TTGTACTATTGACACCGGCTC.

Insert primer (Lbb) (combine with RP): GCGTGGACCGCTTGCTGCAACT. RT-PCR primers used for insertion lines: PILS2 Fw, GCGATCATTTATCGGATC AGT; PILS2 Rev, TTGCATACCTTGACAGTAGTC; PILS5 Fw, TGTTGA AGCCCGTAATCCATGAAC; PILS5 Rev, TTCATTGCGGACCCTTTAAT CAGC.

qPCR primers: PILS1 Fw, CGGTAACACAGCTCCACTTC; PILS1 Rev, GCAACAAGTAACGCACAACC; PILS2 Fw, GTGATGCTTGTACTTGGTGG TATG; PILS2 Rev, AACTTGAACATTGGATCTGCTGAG; PILS3 Fw, AGGCG ACCATGCAAGTGTTG; PILS3 Rev, GTGGTACAGCTAGATGACAGTGAG; PILS4 Fw, TGTCATACTAAGCCTCCTTCAC; PILS4 Rev, CTCGCAACTCTC AGAATCTCC; PILS5 Fw, CTTGGAATAGTCTGTGTTCCGGTAC; PILS5 Rev, GCACTGAGCATTCGCTTGAG; PILS6 Fw, GCCTACATCAGTGCTCTCAG; PILS6 Rev, GCACTGAGCATTCGCTTGAG; PILS7 Fw, TCCTCCAGACCCTC TCTTTCCG; PILS7 Rev, ACAAGAAGATGACCGAGCACTC; EiF4a Fw, CTGGAGGTTTTGAGGCTGGTAT; EiF4a Rev, CCAAGGGTGAAAGCAAG AAGA.

Cloning primers: gDNA/cDNA: PILS1\_Fw, GGGGACAAGTTTGTACAAAA AAGCAGGCTCGATGAGGATGAGGCTTTTGGATC; PILS1\_Rev, GGGGAC CACTTTGTACAAAAAGCTGGGTC(TCA)GGCTACGAGCCACATGAAG AATG; PILS2\_Fw, GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGT CAGGTTTCTCCAGTGGA; PILS2\_Rev, GGGGACCACTTTGTACAAGAAA GCTGGGTC(TCA)TTGCATACCTTGACAGTAGTCTC; PILS3\_Fw, GGGG ACAAGTTTGTACAAAAAAGCAGGCTCGATGGTGAAGCTTTTGGAGCTG

TTT; PILS3\_Rev, GGGGACCACTTTTGTACAAGAAAGCTGGGTC(TCA)AGC TACAAGCCACATGAAGAATG; PILS4\_Fw, GGGGACAAGTTTGTACAAAA AAGCAGGCTCGATGAAGCTTTTGGAGTTGTTC; PILS4\_Rev, GGGGAC CACTTTGTACAAGAAAGCTGGGTC(TCA)TGTCACAAGCCACATGAAGA ATG; PILS5\_Fw, GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGG ATTCTGGTCTGTTGTTGGA; PILS5\_Rev, GGGGACCACTTTTGTACAAGAAA GCTGGGTC(TCA)GACTAACAAGTGAAGGAAGATGG; PILS6\_Fw, GGGGA CAAGTTTGTACAAAAAAGCAGGCTCGATGATTGCTCGGATCCTTGCCG; PILS6\_Rev, GGGGACCACTTTTGTACAAGAAAGCTGGGTC(TCA)GAAGAG TATGTTAATGTAGAGTAC; PILS7\_Fw, GGGGACAAGTTTGTACAAAAA GCAGGCTCGATGGGTTTCTTAGAGTTGTTGGA; PILS7\_Rev, GGGGACCA CTTTGTACAAGAAAGCTGGGTC(TCA)GGAGAGGATGGAGAGGAAGAT GG.

Promoter: PILS2\_Fw, GGGGACAAGTTTGTATAGAAAAAGTTGCGAACTCC ATTGTTAACAGTAATAGC; PILS2\_Rev, GGGGACTGCTTTTTTGTACAAA CTTGCTCGATCTCACTATGTAAAGCTCG; PILS5\_Fw, GGGGACAAGTT TGTATAGAAAAAGTTGCGCAATATACGTGACGTGGTCCACT; PILS5\_Rev, GGGGACTGCTTTTTTGTACAAAAGTTGCTTTTATGTGGTTCTTTAGAC. Destination vectors: *pPILS::GFP/GUS::pKGWFS7*, *p35S::PILS::GFP::pH7FWG2.0*, *p35S::GFP::PILS::pH7WGF2.0*, *p35S::PILS::RFP::pK7RWG2.0*, *p35S::RFP::PILS::pK7WGR2.0* and *p35S::PILS\_D::pH7WGD2.1* (ref. 33), and *pPILS::PILS::GFP::pK7m34GW.0* (ref. 34).

Oestradiol-inducible PILS: *pMDC7\_B(pUBQ)*<sup>35</sup> (p35S promoter was exchanged by the pUBQ10 promoter).

22. Sauer, M., Paciorek, T., Benkova, E. & Friml, J. Immunocytochemical techniques for whole-mount *in situ* protein localization in plants. *Nature Protocols* **1**, 98–103 (2006).
23. Delbarre, A., Muller, P., Imhoff, V. & Guern, J. Comparison of mechanisms controlling uptake and accumulation of 2,4-dichlorophenoxy acetic acid, naphthalene-1-acetic acid, and indole-3-acetic acid in suspension-cultured tobacco cells. *Planta* **198**, 532–541 (1996).
24. Bailly, A. *et al.* Modulation of P-glycoproteins by auxin transport inhibitors is mediated by interaction with immunophilins. *J. Biol. Chem.* **283**, 21817–21826 (2008).
25. Dobrev, P. I., Havlíček, L., Vágner, M., Malbeck, J. & Kamínek, M. Purification and determination of plant hormones auxin and abscisic acid using solid phase extraction and two-dimensional high performance liquid chromatography. *J. Chromatogr. A* **1075**, 159–166 (2005).
26. Dobrev, P. I. & Kamínek, M. Fast and efficient separation of cytokinins from auxin and abscisic acid and their purification using mixed-mode solid-phase extraction. *J. Chromatogr. A* **950**, 21–29 (2002).
27. Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl Acad. Sci. USA* **95**, 5857–5864 (1998).
28. Letunic, I., Doerks, T. & Bork, P. SMART 6: recent updates and new developments. *Nucleic Acids Res.* **37**, 229–232 (2009).
29. Tusnady, G. E. & Simon, I. The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849–850 (2001).
30. Spyropoulos, I. C., Liakopoulos, T. D., Bagos, P. G. & Hamodrakas, S. J. TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics* **20**, 3258–3260 (2004).
31. Proost, S. *et al.* PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell* **21**, 3718–3731 (2009).
32. Pěnčík, A. *et al.* Isolation of novel indole-3-acetic acid conjugates by immunoaffinity extraction. *Talanta* **80**, 651–655 (2009).
33. Karimi, M., De Meyer, B. & Hilson, P. Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103–105 (2005).
34. Karimi, M., Inze, D. & Depicker, A. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7**, 193–195 (2002).
35. Curtis, M. D. & Grossniklaus, U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462–469 (2003).



# Regulation of circadian behaviour and metabolism by REV-ERB- $\alpha$ and REV-ERB- $\beta$

Han Cho<sup>1</sup>, Xuan Zhao<sup>1</sup>, Megumi Hatori<sup>2</sup>, Ruth T. Yu<sup>1</sup>, Grant D. Barish<sup>1</sup>, Michael T. Lam<sup>3</sup>, Ling-Wa Chong<sup>1</sup>, Luciano DiTacchio<sup>2</sup>, Annette R. Atkins<sup>1</sup>, Christopher K. Glass<sup>3</sup>, Christopher Liddle<sup>4</sup>, Johan Auwerx<sup>5</sup>, Michael Downes<sup>1</sup>, Satchidananda Panda<sup>2</sup> & Ronald M. Evans<sup>1,6</sup>

The circadian clock acts at the genomic level to coordinate internal behavioural and physiological rhythms via the CLOCK-BMAL1 transcriptional heterodimer. Although the nuclear receptors REV-ERB- $\alpha$  and REV-ERB- $\beta$  have been proposed to form an accessory feedback loop that contributes to clock function<sup>1,2</sup>, their precise roles and importance remain unresolved. To establish their regulatory potential, we determined the genome-wide *cis*-acting targets (cistromes) of both REV-ERB isoforms in murine liver, which revealed shared recognition at over 50% of their total DNA binding sites and extensive overlap with the master circadian regulator BMAL1. Although REV-ERB- $\alpha$  has been shown to regulate *Bmal1* expression directly<sup>1,2</sup>, our cistromic analysis reveals a more profound connection between BMAL1 and the REV-ERB- $\alpha$  and REV-ERB- $\beta$  genomic regulatory circuits than was previously suspected. Genes within the intersection of the BMAL1, REV-ERB- $\alpha$  and REV-ERB- $\beta$  cistromes are highly enriched for both clock and metabolic functions. As predicted by the cistromic analysis, dual depletion of *Rev-erb- $\alpha$*  and *Rev-erb- $\beta$*  function by creating double-knockout mice profoundly disrupted circadian expression of core circadian clock and lipid homeostatic gene networks. As a result, double-knockout mice show markedly altered circadian wheel-running behaviour and deregulated lipid metabolism. These data now unite REV-ERB- $\alpha$  and REV-ERB- $\beta$  with PER, CRY and other components of the principal feedback loop that drives circadian expression and indicate a more integral mechanism for the coordination of circadian rhythm and metabolism.

The circadian clock is a transcriptional mechanism that coordinates both behavioural and physiological processes such as the sleep–wake cycle and food intake. The existing model for the mammalian core molecular clock involves a transcriptional negative-feedback loop in which the transactivation of E-box-containing genes by CLOCK and BMAL1 is inhibited by the expression of *Per1* and *Cry1*, themselves under transcriptional control of E boxes<sup>3</sup>. As functional redundancy is common for core clock components, deletion of multiple paralogues in mice is often required to uncover relevant phenotypes such as perturbations in circadian gene expression, metabolism and wheel-running behaviour<sup>4–7</sup>.

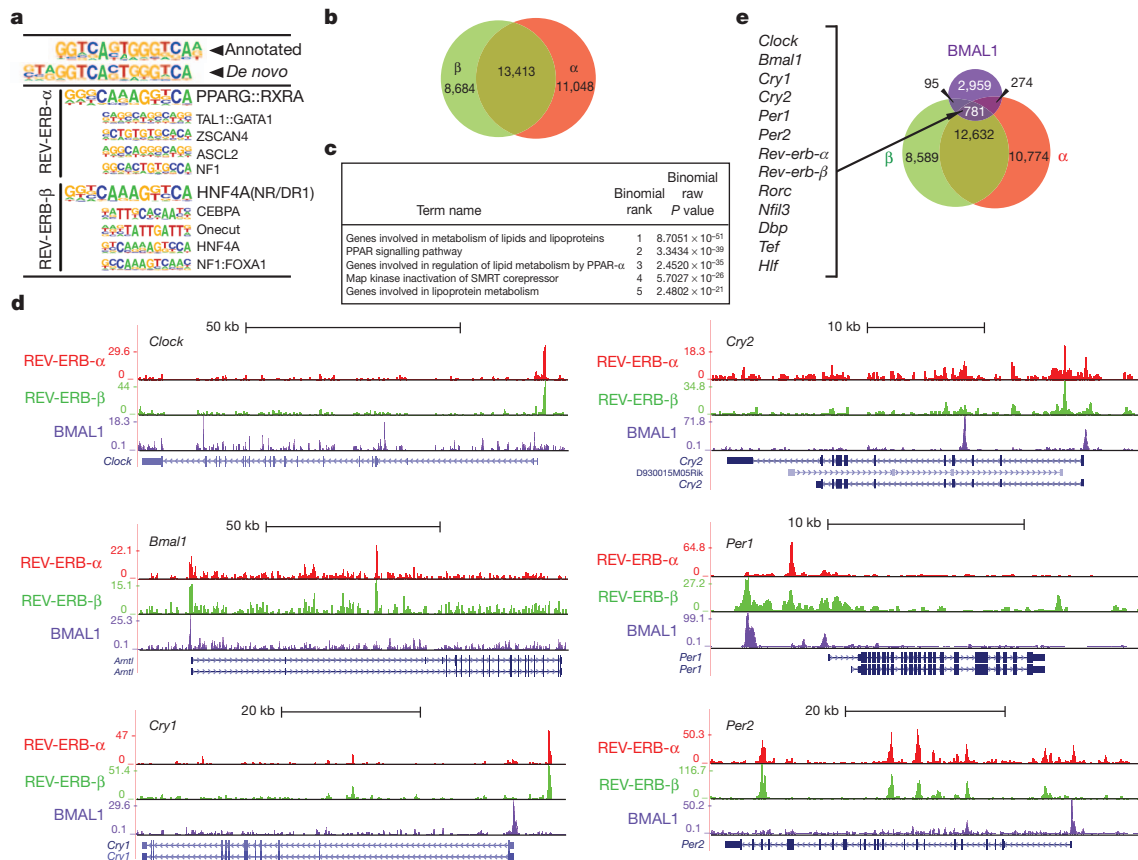
As REV-ERB- $\alpha$ , REV-ERB- $\beta$ , ROR- $\alpha$ , ROR- $\beta$  and ROR- $\gamma$  bind to a common response element (the RORE), their intrinsic repressive and inductive activities, respectively, are believed to establish the rhythmic expression of target genes such as *Bmal1*. However, the partial penetrance and mild period phenotype of *Rev-erb- $\alpha$* <sup>−/−</sup> mice<sup>1</sup> (essentially intact circadian wheel-running behaviour) and the modest rhythmic phenotype upon partial *Rev-erb- $\beta$*  depletion of *Rev-erb- $\alpha$* <sup>−/−</sup> cultured cells<sup>2</sup>, suggested that they are not required for principal core clock function. Rather, REV-ERBs are proposed to form an accessory feedback loop that stabilizes the clock and has a role in receiving input signals or transmitting output pathways<sup>2,8</sup>.

To clarify the regulatory potential of REV-ERB- $\alpha$  and REV-ERB- $\beta$  in circadian regulation, we generated isoform-specific antibodies (Supplementary Methods and Supplementary Fig. 1) and determined their genome-wide *cis*-acting targets (cistromes) in the liver at Zeitgeber time (ZT) 8, the peak of their protein expression (data not shown and ref. 9). *De novo* motif analysis (Fig. 1a) revealed that *in vivo*, in addition to the classic REV-ERB direct repeat 2 (DR2) motifs, other nuclear receptor binding sites (particularly DR1) are predominant in the DNA elements bound by both REV-ERB- $\alpha$  and REV-ERB- $\beta$ <sup>10</sup>. Although not surprising, extensive overlap was observed between the REV-ERB cistromes, with commonly bound peaks accounting for 54.8% and 60.7% of the total REV-ERB- $\alpha$  and REV-ERB- $\beta$  peaks, respectively (Fig. 1b). The limited overlap of our REV-ERB- $\alpha$  cistrome with one that has been published previously<sup>11</sup> (<50%) can most likely be attributed to differences in the antibody specificities (Supplementary Figs 2 and 3). Pathway analyses of our REV-ERB- $\alpha$  and REV-ERB- $\beta$  overlapping peaks<sup>12</sup> revealed an enrichment in lipid metabolism genes (Fig. 1c), consistent with the hyperlipidaemic phenotype previously observed in *Rev-erb- $\alpha$*  null mice<sup>13</sup>. Notably, loci encoding circadian clock genes (*Clk*, *Bmal1*, *Cry1*, *Cry2*, *Per1*, *Per2*; see Fig. 1d) were also enriched in the REV-ERB- $\alpha$ / $\beta$  cistromic overlap, suggesting that the coordinated actions of both REV-ERB isoforms are directly linked with clock function. A comparison of the REV-ERB- $\alpha$ / $\beta$  cistrome with published BMAL1 binding sites<sup>14</sup> revealed that 28% of BMAL1 peaks (at ZT6 and ZT10) were shared with the REV-ERB- $\alpha$ / $\beta$  (ZT8) cistrome and 68% of these peaks (781) were occupied by all three transcription factors. Clear binding sites for each of these transcription factors were found on ‘core clock’ gene loci as well as on many clock-controlled target genes (*Rev-erb- $\alpha$* , *Rev-erb- $\beta$* , *Ror*, *Dbp*, *Hlf*, *Tef*, *Nfil3*; see Fig. 1d, e). In addition to circadian annotated loci, the BMAL1, REV-ERB- $\alpha$ , REV-ERB- $\beta$  triple intersection is highly enriched for genes in the receptor tyrosine kinase signalling pathway as well as those known for energy homeostasis (Supplementary Table 1). Their confluence at hundreds of clock and clock output genes indicates that beyond a simple ‘binary relationship’, REV-ERB- $\alpha$ / $\beta$  and BMAL1 cooperate to regulate clock and clock output genes coordinately.

To test the above proposal, we used Cre/lox recombination to generate three genetically modified mouse lines in C57BL/6 backgrounds harbouring global, tissue-specific, or conditional knockouts of *Rev-erb- $\alpha$*  and *Rev-erb- $\beta$*  loci. Global *Rev-erb- $\alpha$* <sup>−/−</sup> and *Rev-erb- $\beta$* <sup>−/−</sup> single knockout mice were generated using a *CMV-cre* transgenic allele (detailed in Supplementary Methods). Homozygous deletion of *Rev-erb- $\beta$*  did not cause overt gross abnormalities, lethality or infertility, whereas *Rev-erb- $\alpha$* <sup>−/−</sup> mice, although viable, show frequent postnatal lethality (before 2 weeks of age) (Supplementary Table 2), with survivors exhibiting diminished fertility in both sexes. In contrast,

<sup>1</sup>Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, California 92037, USA. <sup>2</sup>Regulatory Biology Laboratory, Salk Institute for Biological Studies, La Jolla, California 92037, USA.

<sup>3</sup>Department of Cellular and Molecular Medicine, School of Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0651, USA. <sup>4</sup>The Storr Liver Unit, Westmead Millennium Institute and University of Sydney, Westmead Hospital, Westmead, New South Wales 2145, Australia. <sup>5</sup>Ecole Polytechnique Fédérale in Lausanne CH-1015, Lausanne, Switzerland. <sup>6</sup>Howard Hughes Medical Institute, Salk Institute for Biological Studies, La Jolla, California 92037, USA.



**Figure 1 | Cistromic analyses of REV-ERB- $\alpha$  and REV-ERB- $\beta$  in liver. a**, *De novo* HOMER motif analysis of *in vivo* REV-ERB- $\alpha$  and REV-ERB- $\beta$  binding. **b**, Venn diagram depicting the unique and common REV-ERB- $\alpha$  and REV-ERB- $\beta$  bound peaks. **c**, Commonly bound REV-ERB- $\alpha$  and REV-ERB- $\beta$  peaks are enriched for genes involved in lipid metabolism and associated with PPARs.

*Rev-erb- $\alpha$* <sup>-/-</sup> mice on a mixed 129/Sv and C57BL/6 background were reported to have reduced fertility but no postnatal lethality<sup>15</sup>.

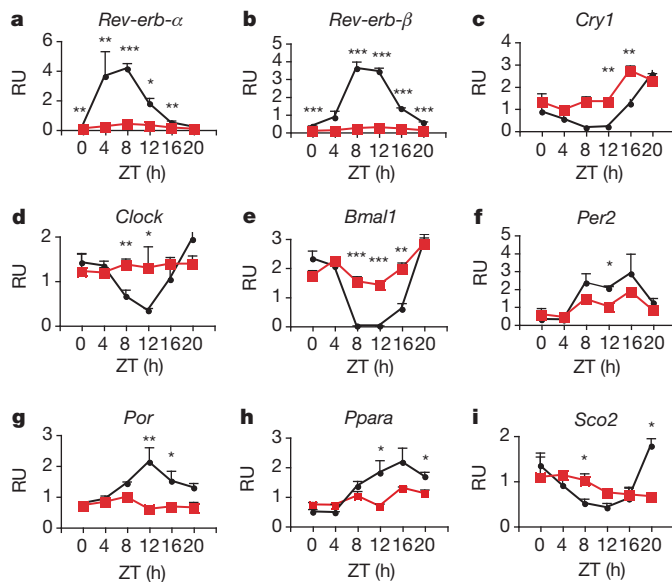
Liver-specific *Rev-erb- $\alpha$* <sup>-/-</sup> *Rev-erb- $\beta$* <sup>-/-</sup> mice were generated to assess combinatorial REV-ERB signalling in hepatic circadian gene expression via an albumin-Cre transgene<sup>16</sup>. These liver-specific double knockout mice (*Rev-erb- $\alpha$* <sup>lox/lox</sup> *Rev-erb- $\beta$* <sup>lox/lox</sup> albumin-Cre (L-DKO)) were born at the expected Mendelian ratio, effectively bypassing the frequent lethality of the *Rev-erb- $\alpha$* <sup>-/-</sup> global deletion. Depletion of both *Rev-erb- $\alpha$*  and *Rev-erb- $\beta$*  (Fig. 2a, b) disrupted the circadian, and often total, expression of many hepatic core clock genes (Fig. 2c–f) and presumed output genes (Fig. 2g–i).

Microarray analysis of gene expression over a 24-h time course (presented as Zeitgeber time) revealed massive differences between the wild-type and L-DKO liver. Using CircWave v3.3 (ref. 17) software, we determined that of the ~900 genes showing circadian expression in wild-type liver, the rhythmicity of more than 90% was perturbed in the L-DKO liver (Fig. 3a). The severity of the circadian disruption was comparable to that described for mice deficient in core clock components<sup>18,19</sup>, emphasizing the functional impact associated with the extensive intersection between REV-ERB- $\alpha/\beta$  and BMAL1 cistromes. Furthermore, the gene ontology of the REV-ERB- $\alpha/\beta$ -dependent circadian transcriptome (those genes that lose rhythm in L-DKO mice; Fig. 3b) also mirrors that of their cistromes (Fig. 1c and Supplementary Table 1). This is of interest as the disrupted oscillation of these genes in the liver occurred *in vivo* in the presence of otherwise wild-type entraining signals and fully intact extra-hepatic clockwork (Supplementary Fig. 5). Notably, residual rhythmic expression of some genes was maintained in the L-DKO mice, indicating that they may be controlled by systemic cues, as suggested previously<sup>20</sup>.

**d**, REV-ERB- $\alpha$ , REV-ERB- $\beta$  and BMAL1 binding at canonical circadian clock genes. The left axis indicates sequence-tag counts. **e**, BMAL1 cistrome significantly overlaps with REV-ERB- $\alpha$  and REV-ERB- $\beta$ . Examples of clock-related genes in overlap are listed, and selected peaks are shown in Supplementary Fig. 4.

To link the cistromic and transcriptomic data functionally, we examined the gene expression levels of REV-ERB/BMAL1 common cistromic sites. A comparison of the liver expression levels across ZT8 to ZT16 revealed that 45% of the co-occupied genes were significantly perturbed in the liver-specific DKO compared with wild-type liver (Supplementary Table 3). This strong correlation between REV-ERB occupancy and perturbed gene expression corroborates a direct role for REV-ERBs in maintaining rhythmic expression patterns for many of the above genes.

The benchmark assay for circadian dysfunction is wheel-running behaviour, where the persistence or fragmentation of running rhythm is assessed in constant darkness after prior entrainment to a 12-h light/dark cycle. Actograms recorded for *Rev-erb- $\alpha$* <sup>-/-</sup> mice in a pure C57BL/6 background displayed a fully penetrant period shortening of 30 min (Fig. 4a, b), contrasting with the lack of fully penetrant phenotype reported in mixed background studies<sup>1,15</sup>. No changes in the activity levels, activity consolidation, or phase angle of entrainment (activity onset relative to dark onset) were observed for *Rev-erb- $\alpha$* <sup>-/-</sup> mice. *Rev-erb- $\beta$* <sup>-/-</sup> mice showed no change in any of the parameters of circadian activity rhythms (Fig. 4a, b), raising the possibility that the circadian clock function is independent of REV-ERB- $\beta$ . Given the functional redundancy demonstrated for the established components of the core clock, we sought to address whether REV-ERB- $\alpha$  and REV-ERB- $\beta$  might cooperatively influence circadian rhythms in wheel running. Using the tamoxifen-activated *creER* transgene, inducible double knockout (iDKO) mice were generated<sup>21</sup>, allowing ubiquitous deletion of both genes in adulthood (see Supplementary Methods). Whereas the wheel-running activity of *Rev-erb- $\alpha$* <sup>lox/lox</sup> *Rev-erb- $\beta$* <sup>lox/lox</sup> mice treated with tamoxifen was normal, the phenotype of mice



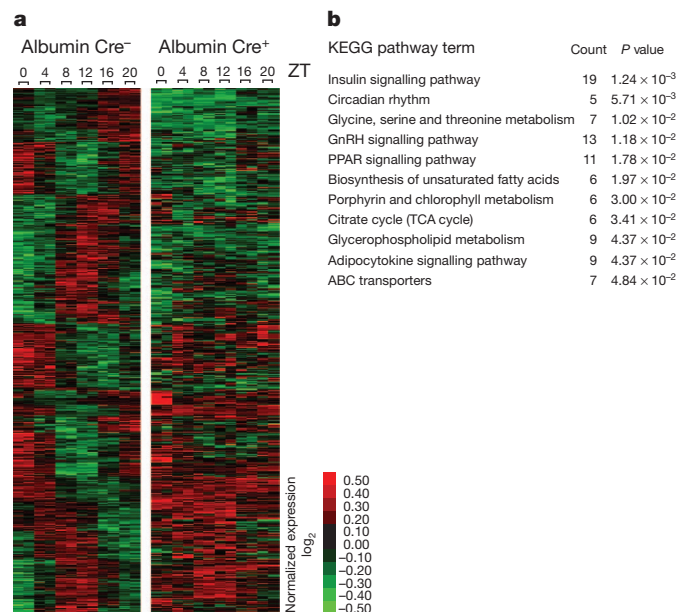
**Figure 2 | Circadian gene expression of many canonical core clock genes and output genes are disrupted in livers of *Rev-erb-α*<sup>lox/lox</sup> *Rev-erb-β*<sup>lox/lox</sup> albumin-Cre (L-DKO) mice.** a–i, The expression levels of *Rev-erb-α* (a), *Rev-erb-β* (b), canonical core clock genes (*Cry1*, *Clock*, *Bmal1* and *Per2*) (c–f) and presumed output genes (*Por*, *Ppara* and *Sco2*) (g–i) in livers from L-DKO (albumin-Cre positive, red labels) and wild-type (albumin-Cre negative, black labels) mice are shown. Livers ( $n = 3$ ) were collected at each indicated ZT under a 12-h light/dark cycle. QPCR was performed in technical triplicates. Relative units (RU) normalized with *36B4*. Error bars indicate standard error of the mean; statistical significance was determined by Student's *t*-test (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ).

bearing the *creER* transgene was more severe than the additive effects of the individual knockouts (Fig. 4a–c). The iDKO mice showed reduced and severely fragmented activity, and advanced phase angle of entrainment, features also found in *Bmal1*<sup>−/−</sup> mice<sup>22</sup>. Furthermore,

as shown in the periodogram in Fig. 4c, in mice with detectable activity rhythms, the free-running period length under constant darkness was shortened by as much as 2.5 h (see also Supplementary Tables 4 and 5). The severe activity defects in iDKO mice is unexpected when one considers the properties of the single knockouts, and supports a strongly selected, novel and cooperative role for REV-ERB- $\alpha$  and REV-ERB- $\beta$  in rhythmic behaviour<sup>4–7</sup>.

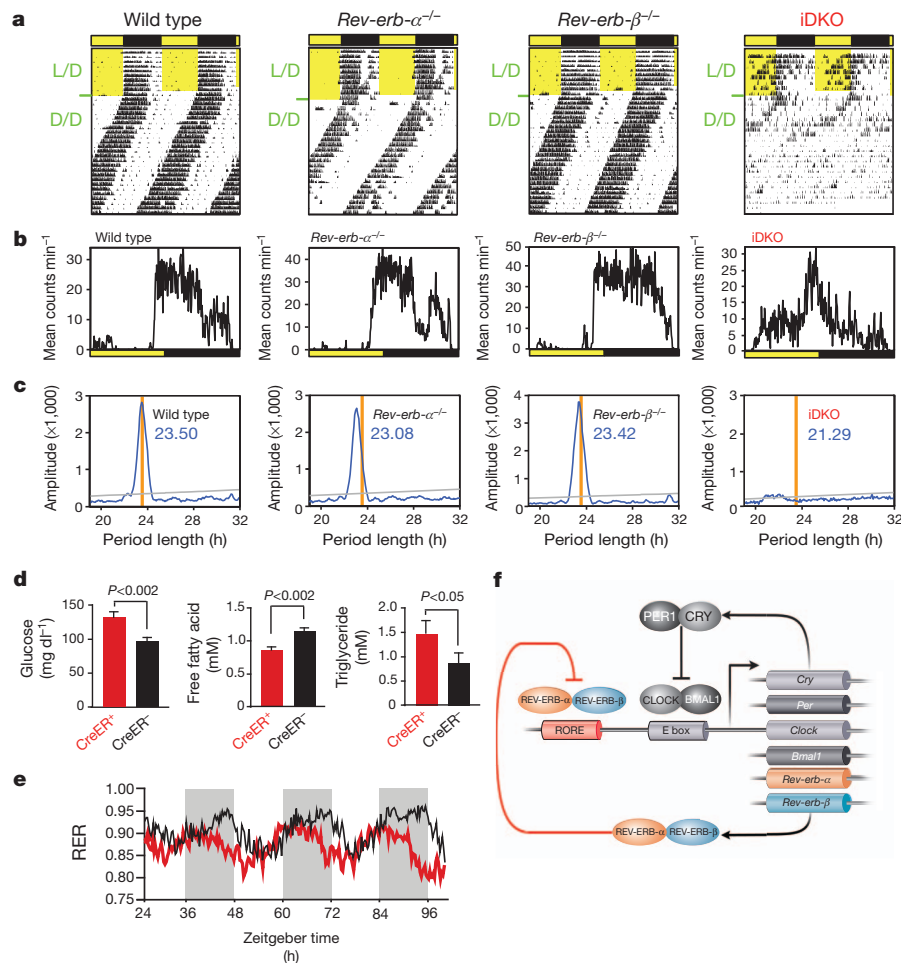
Mice and humans with disrupted circadian rhythms have been shown to develop metabolic disorders including hyperlipidaemia and hyperglycaemia, suggesting a link between proper clock function and metabolism<sup>23</sup>. To investigate if REV-ERB activity is similarly required for normal metabolic regulation, we compared metabolic parameters of tamoxifen-treated control and iDKO mice. Treated iDKO mice displayed increased circulating glucose and triglyceride levels, and a reduction in the level of free fatty acids compared to control littermates (Fig. 4d). The reduced fatty acids may reflect a more oxidative metabolism of the double knockout mice as exemplified by reduced RERs. We also note a distinct RER phase shift in constant dark conditions (Fig. 4e), indicating dysregulation of overall body metabolism. Although previous studies have linked REV-ERB- $\alpha$  with lipid homeostasis<sup>13,24,25</sup>, the presence of both REV-ERB- $\alpha$  and REV-ERB- $\beta$  on multiple key lipid and bile acid regulatory genes including *LXRα* (also called *Nr1h3*), *Fxr* (*Nr1h4*), *Apoc3*, *Cyp7a1*, *SHP* (*Nr0b2*), *Insig2* and SREBP (*Srebf1*) provides a molecular mechanism for rhythmic cholesterol and bile acid metabolism. The current cistromic studies elevate REV-ERB- $\beta$  to equal prominence with REV-ERB- $\alpha$  in the transcriptional regulation of these pathways, and identify additional direct gene targets (Supplementary Fig. 6). The severity of the metabolic phenotype observed in the iDKO mice compared to the REV-ERB- $\alpha$  knockout mice is consistent with the cistromic analyses of the REV-ERBs.

Defining the genetic mechanisms that comprise the circadian clock is central to understanding how genomic rhythms are transformed into behavioural and metabolic physiology. In this study, using a combination of genome-wide cistromic profiling, mRNA expression analysis and three new targeted gene knockout mouse models, we unequivocally demonstrate a more critical and central role for REV-ERB- $\alpha/\beta$  circadian clock regulation than previously suspected. The genomic analysis confirms *Bmal1* as a direct REV-ERB- $\alpha/\beta$  target, while at the genome-wide level we find a predominance of both core clock and circadian output genes targeted by REV-ERBs, indicating a more integral role for this family in the core clock than previously considered. These findings indicate that a dynamic balance between BMAL1 and REV-ERB- $\alpha/\beta$  cistromes is used to regulate both core clock and clock output genes by co-localizing opposing epigenetic regulators to shared genomic targets. *In vivo*, using targeted and inducible double knockout mice, we demonstrate that REV-ERB- $\alpha$  and REV-ERB- $\beta$  together function as integral drivers of the circadian clock, rather than simply as stabilizers of an output, thereby redefining the established paradigm for these receptors. The similarity of the DKO circadian phenotype to compound core clock mutants (*Per1*<sup>−/−</sup> *Per2*<sup>−/−</sup>, *Cry1*<sup>−/−</sup> *Cry2*<sup>−/−</sup>)<sup>5–7</sup> in severity, penetrance and strong period shortening is more reflective of a pacemaker rather than a stabilizer of rhythm. Together, this leads to a model of mutual direct regulation, with BMAL1 controlling one loop and REV-ERB- $\alpha/\beta$  a second loop of the core clock, with both loops using cistromic convergence to coordinate key clock and metabolic functions (Fig. 4f). The adaptive feature of the circadian clock enables its control of sleep–wake cycles, physiological rhythms, energy homeostasis and behaviour. In contrast, disruption of rhythm spawns a range of problems from jet lag, to more profound sleep disorders, obesity, metabolic disease, immune function and cancer<sup>26–29</sup>. As partnered regulators, the recent development of both potent REV-ERB agonists (that enhance repression)<sup>25</sup> and REV-ERB antagonists (that relieve repression)<sup>30</sup> provides a new therapeutic approach to both reset disrupted rhythms and re-establish metabolic balance.



**Figure 3 | Broad disruption of circadian transcriptome in the absence of *Rev-erb-α* and *Rev-erb-β*.** a, Heat map of genes with circadian expression in wild-type (left panel) and L-DKO (right panel) livers. A total of 1,227 unique accession numbers were selected based on false discovery rate  $< 0.05$ . b, Genes expressed in a circadian manner that lose rhythm are highly associated with circadian and energy homeostasis functions as assessed by KEGG pathway analysis.





**Figure 4 | Loss of both *Rev-erb-α* and *Rev-erb-β* results in disrupted circadian wheel-running behaviour and metabolic shift.** **a–c**, Voluntary locomotor activity of wild-type, *Rev-erb-α*<sup>-/-</sup>, *Rev-erb-β*<sup>-/-</sup> and *Rev-erb-α*<sup>-/-</sup>*Rev-erb-β*<sup>-/-</sup> (iDKO) mice. **a**, Actograms showing wheel-running activity in constant darkness after prior entrainment in light/dark. **b**, Activity profiles during light/dark cycles. **c**, Chi-squared periodogram of the initial 20 days in constant darkness (*n* = 5–9 for each mutant strain; *n* = 5–6 littermate

## METHODS SUMMARY

For chromatin immunoprecipitation (ChIP), 5-month-old male C57BL/6J mice were killed by CO<sub>2</sub> asphyxiation at ZT8. Livers were removed and pooled for ChIP, processing and sequencing. Tamoxifen-induction of Cre recombinase activity was accomplished by daily intraperitoneal injection of 2 mg tamoxifen (Sigma) in 100 μl corn oil (Sigma) for 7 days. Three-month-old tamoxifen-treated animals were subjected to wheel-running assays for 7 days after the end of treatment. For gene expression analysis, 5–6-month-old males were killed by cervical dislocation at indicated ZT points for rapid dissection and snap freezing of the tissues. During the dark cycle, procedures were performed under red light. Detailed methods are provided in the Supplementary Information.

Received 14 February; accepted 21 March 2012.

Published online 29 March 2012.

1. Preitner, N. *et al.* The orphan nuclear receptor REV-ERB $\alpha$  controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251–260 (2002).
2. Liu, A. C. *et al.* Redundant function of REV-ERB $\alpha$  and  $\beta$  and non-essential role for Bmal1 cycling in transcriptional regulation of intracellular circadian rhythms. *PLoS Genet.* **4**, e1000023 (2008).
3. Gekakis, N. *et al.* Role of the CLOCK protein in the mammalian circadian mechanism. *Science* **280**, 1564–1569 (1998).
4. DeBruyne, J. P., Weaver, D. R. & Reppert, S. M. CLOCK and NPAS2 have overlapping roles in the suprachiasmatic circadian clock. *Nature Neurosci.* **10**, 543–545 (2007).
5. Zheng, B. *et al.* Nonredundant roles of the *mPer1* and *mPer2* genes in the mammalian circadian clock. *Cell* **105**, 683–694 (2001).

controls). Representative actograms from individual mice are shown. The period length of wild-type mice is shown as a vertical orange line. **d**, Fasting glucose (*n* = 6), free fatty acid (*n* = 6) and triglyceride (*n* = 6) levels in iDKO and wild-type mice. **e**, Respiratory exchange ratio (RER,  $v_{CO_2}/v_{O_2}$ ) for wild-type (black) and iDKO (red) mice (*n* = 4). **f**, Model depicting the activating (CLOCK-BMAL1) and repressive (REV-ERB- $\alpha$ -REV-ERB- $\beta$ ) transcriptional complexes, the coordinate actions of which generate rhythmic gene expression.

6. van der Horst, G. T. *et al.* Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* **398**, 627–630 (1999).
7. Vitaterna, M. H. *et al.* Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2. *Proc. Natl Acad. Sci. USA* **96**, 12114–12119 (1999).
8. Levi, F. & Schibler, U. Circadian rhythms: mechanisms and therapeutic implications. *Annu. Rev. Pharmacol. Toxicol.* **47**, 593–628 (2007).
9. Ukai-Tadenuma, M. *et al.* Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell* **144**, 268–281 (2011).
10. Yang, X. *et al.* Nuclear receptor expression links the circadian clock to metabolism. *Cell* **126**, 801–810 (2006).
11. Feng, D. *et al.* A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. *Science* **331**, 1315–1319 (2011).
12. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
13. Le Martelot, G. *et al.* REV-ERB $\alpha$  participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biol.* **7**, e1000181 (2009).
14. Rey, G. *et al.* Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biol.* **9**, e1000595 (2011).
15. Chomez, P. *et al.* Increased cell death and delayed development in the cerebellum of mice lacking the rev-erbA $\alpha$  orphan receptor. *Development* **127**, 1489–1498 (2000).
16. Postic, C. *et al.* Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic beta cell-specific gene knock-outs using Cre recombinase. *J. Biol. Chem.* **274**, 305–315 (1999).
17. Oster, H., Damerow, S., Hut, R. A. & Eichele, G. Transcriptional profiling in the adrenal gland reveals circadian regulation of hormone biosynthesis genes and nucleosome assembly genes. *J. Biol. Rhythms* **21**, 350–361 (2006).

18. Miller, B. H. *et al.* Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. *Proc. Natl Acad. Sci. USA* **104**, 3342–3347 (2007).
19. Hatanaka, F. *et al.* Genome-wide profiling of the core clock protein BMAL1 targets reveals a strict relationship with metabolism. *Mol. Cell. Biol.* **30**, 5636–5648 (2010).
20. Kornmann, B. *et al.* System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS Biol.* **5**, e34 (2007).
21. Hayashi, S. & McMahon, A. P. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* **244**, 305–318 (2002).
22. Bunge, M. K. *et al.* Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **103**, 1009–1017 (2000).
23. Huang, W. *et al.* Circadian rhythms, sleep, and metabolism. *J. Clin. Invest.* **121**, 2133–2141 (2011).
24. Raspé, E. *et al.* Identification of Rev-erb $\alpha$  as a physiological repressor of apoC-III gene transcription. *J. Lipid Res.* **43**, 2172–2179 (2002).
25. Kumar, N. *et al.* Regulation of adipogenesis by natural and synthetic REV-ERB ligands. *Endocrinology* **151**, 3015–3025 (2010).
26. Gibbs, J. E. *et al.* The nuclear receptor REV-ERB $\alpha$  mediates circadian regulation of innate immunity through selective regulation of inflammatory cytokines. *Proc. Natl Acad. Sci. USA* **109**, 582–587 (2012).
27. Huang, W., Ramsey, K. M., Marcheva, B. & Bass, J. Circadian rhythms, sleep, and metabolism. *J. Clin. Invest.* **121**, 2133–2141 (2011).
28. Lamia, K. A. *et al.* AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. *Science* **326**, 437–440 (2009).
29. Lamia, K. A. *et al.* Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature* **480**, 552–556 (2011).
30. Kojetin, D., Wang, Y., Kamenecka, T. M. & Burris, T. P. Identification of SR8278, a synthetic antagonist of the nuclear heme receptor REV-ERB. *ACS Chem. Biol.* **6**, 131–134 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank S. Kaufman, J. Alvarez, E. Banayo, H. Juguilon, S. Jacinto and H. Le for technical assistance; and L. Ong and S. Ganley for administrative assistance. We also thank L. Pei for discussion. R.M.E. is an Investigator of the Howard Hughes Medical Institute at The Salk Institute for Biological Studies and March of Dimes Chair in Molecular and Developmental Biology. H.C. is a recipient of National Research Service Award (T32-HL007770). This work was supported by National Institutes of Health Grants (DK062434, DK057978, DK090962, DK091618 and HL105278), National Health and Medical Research Council of Australia Project Grants (NHMRC 512354 and 632886), the Helmsley Charitable Trust, the Glenn Foundation and the Howard Hughes Medical Institute.

**Author Contributions** X.Z., G.D.B., R.T.Y., M.D. and C.L. performed and/or analysed the results from ChIP-seq. R.T.Y., M.D., M.T.L. and C.K.G. performed and/or analysed the results from the microarray experiment. M.H., L.D. and S.P. performed and/or analysed the wheel-running assay and the real-time luciferase assay. J.A. performed gene targeting. H.C. and L.-W.C. performed all experiments. H.C. and R.M.E. designed all experiments, analysed all results and H.C., R.T.Y., M.D., A.R.A., S.P. and R.M.E. wrote the manuscript.

**Author Information** Microarray and ChIP-seq data sets have been deposited in the NCBI Gene Expression Omnibus with the accession number GSE34020. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to R.M.E. ([evans@salk.edu](mailto:evans@salk.edu)).

# Inositol-1,4,5-trisphosphate receptor regulates hepatic gluconeogenesis in fasting and diabetes

Yiguo Wang<sup>1</sup>, Gang Li<sup>2</sup>, Jason Goode<sup>1</sup>, Jose C. Paz<sup>1</sup>, Kunfu Ouyang<sup>3</sup>, Robert Screaton<sup>4,5,6</sup>, Wolfgang H. Fischer<sup>1</sup>, Ju Chen<sup>3</sup>, Ira Tabas<sup>2,7,8</sup> & Marc Montminy<sup>1</sup>

**In the fasted state, increases in circulating glucagon promote hepatic glucose production through induction of the gluconeogenic program. Triggering of the cyclic AMP pathway increases gluconeogenic gene expression via the de-phosphorylation of the CREB co-activator CRTC2 (ref. 1). Glucagon promotes CRTC2 dephosphorylation in part through the protein kinase A (PKA)-mediated inhibition of the CRTC2 kinase SIK2. A number of Ser/Thr phosphatases seem to be capable of dephosphorylating CRTC2 (refs 2, 3), but the mechanisms by which hormonal cues regulate these enzymes remain unclear. Here we show in mice that glucagon stimulates CRTC2 dephosphorylation in hepatocytes by mobilizing intracellular calcium stores and activating the calcium/calmodulin-dependent Ser/Thr-phosphatase calcineurin (also known as PP3CA). Glucagon increased cytosolic calcium concentration through the PKA-mediated phosphorylation of inositol-1,4,5-trisphosphate receptors (InsP<sub>3</sub>Rs), which associate with CRTC2. After their activation, InsP<sub>3</sub>Rs enhanced gluconeogenic gene expression by promoting the calcineurin-mediated dephosphorylation of CRTC2. During feeding, increases in insulin signalling reduced CRTC2 activity via the AKT-mediated inactivation of InsP<sub>3</sub>Rs. InsP<sub>3</sub>R activity was increased in diabetes, leading to upregulation of the gluconeogenic program. As hepatic downregulation of InsP<sub>3</sub>Rs and calcineurin improved circulating glucose levels in insulin resistance, these results demonstrate how interactions between cAMP and calcium pathways at the level of the InsP<sub>3</sub>R modulate hepatic glucose production under fasting conditions and in diabetes.**

We tested a series of Ser/Thr protein phosphatase inhibitors for their ability to block CRTC2 activation in response to glucagon. Exposure to the calcineurin inhibitor cyclosporine A (CsA) disrupted the glucagon-induced dephosphorylation and nuclear translocation of CRTC2, but okadaic acid, an inhibitor of PP1, PP2A and PP4 did not (Fig. 1a and Supplementary Fig. 1a). CsA and other calcineurin inhibitors also reduced cAMP response element (CRE)-luciferase (Luc) reporter activity (Fig. 1a and Supplementary Fig. 1b), but they had no effect in cells expressing phosphorylation-defective (Ser 171, 275 Ala) and therefore active forms of CRTC2 (Supplementary Fig. 1c–e).

On the basis of the ability of CsA to interfere with CRTC2 activation, we considered that calcineurin may promote the dephosphorylation of CRTC2 in response to glucagon. Supporting this idea, CRTC2 contains two consensus (PXIXIT) motifs that mediate an association with calcineurin<sup>3,4</sup> (Supplementary Fig. 2a, b). Moreover, mutation of both motifs disrupted the glucagon-dependent dephosphorylation of CRTC2 (Fig. 1b) and prevented its nuclear translocation (Supplementary Fig. 2c), thereby down-regulating CRE-Luc activation (Fig. 1b).

On the basis of these results, we tested whether calcineurin modulates expression of the gluconeogenic program. Adenoviral overexpression of the calcineurin catalytic subunit in hepatocytes augmented CRTC2 dephosphorylation, CRE-Luc activity, and glucose secretion in

response to glucagon, whereas calcineurin knockdown had the opposite effect (Fig. 1c). Although calcineurin could, in principle, modulate CRTC2 activity indirectly through effects on cAMP signalling, calcineurin overexpression or knockdown did not alter the phosphorylation of cellular PKA substrates in cells exposed to glucagon (Supplementary Fig. 2d).

We examined next whether calcineurin modulates hepatic gluconeogenesis *in vivo*. Modest (twofold) overexpression of calcineurin in liver increased gluconeogenic gene expression, hepatic CRE-Luc activity, and fasting blood glucose concentrations (Fig. 1d and Supplementary Fig. 3a). Conversely, knockdown of hepatic calcineurin reduced expression of the gluconeogenic program and lowered circulating glucose levels (Fig. 1d and Supplementary Fig. 3b), demonstrating that this phosphatase contributes to fasting adaptation in the liver. Calcineurin seemed to stimulate gluconeogenesis via the CREB pathway; depletion of CRTC2 blocked the effects of calcineurin overexpression (Supplementary Fig. 4).

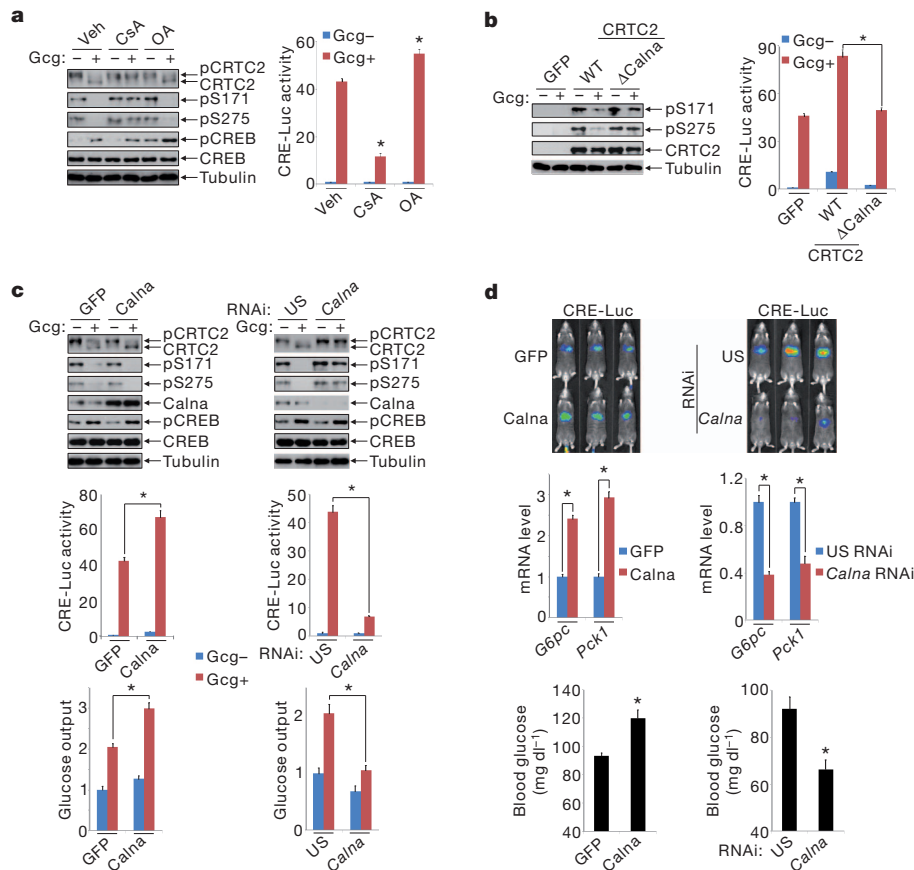
Realizing that calcineurin activity is dependent on increases in intracellular calcium, we tested whether the cAMP pathway stimulates calcium mobilization. Exposure of primary hepatocytes to glucagon triggered a rapid increase in cellular free calcium (Fig. 2a and Supplementary Fig. 5a); these effects were partially reversed by co-treatment with the PKA inhibitor H89 (Supplementary Fig. 5b). The rise in intracellular calcium seems to be critical for CRTC2 activation, because co-incubation with the calcium chelator BAPTA disrupted CRTC2 dephosphorylation and CRE-Luc activation in response to glucagon (Fig. 2b). Arguing against an effect of calcium on cAMP signalling, exposure to BAPTA did not block the PKA-mediated phosphorylation of CREB in response to glucagon.

We imagined that cAMP may increase calcium mobilization through the PKA-dependent phosphorylation of an intracellular calcium channel. In mass spectrometry studies to identify proteins that undergo phosphorylation by PKA in response to glucagon, we recovered the inositol 1,4,5-trisphosphate receptor 1 (InsP<sub>3</sub>R1) from immunoprecipitates of phospho-PKA substrate antiserum (Supplementary Fig. 5c). InsP<sub>3</sub>R1 and its related family members (InsP<sub>3</sub>R2, InsP<sub>3</sub>R3) are calcium release channels that promote the mobilization of endoplasmic reticulum calcium stores following their activation in response to extracellular signals<sup>5–9</sup>. Moreover, cAMP agonists have also been shown to enhance InsP<sub>3</sub>R receptor activity through PKA-mediated phosphorylation.

Inhibiting InsP<sub>3</sub>Rs, either by exposure of hepatocytes to xestospongin C or by knockdown of all three InsP<sub>3</sub>Rs, disrupted cytosolic calcium mobilization and calcineurin activation in response to glucagon and forskolin (Fig. 2a and Supplementary Fig. 6a). Moreover, xestospongin C treatment and InsP<sub>3</sub>R knockdown also blocked the effects of glucagon on CRTC2 dephosphorylation, CRE-Luc activation, and induction of the gluconeogenic program (Fig. 2c and Supplementary Fig. 6a, b). We

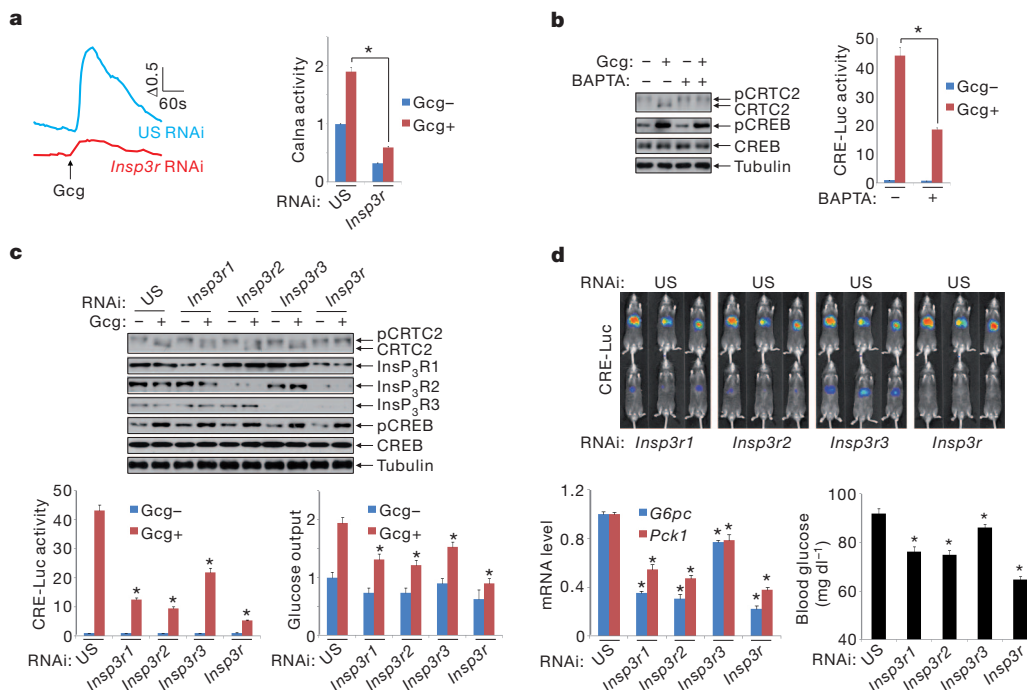
<sup>1</sup>Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. <sup>2</sup>Department of Medicine, Columbia University, New York, New York 10032, USA. <sup>3</sup>Department of Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA. <sup>4</sup>Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. <sup>5</sup>Department of Pediatrics, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. <sup>6</sup>Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario K1H 8L1, Canada. <sup>7</sup>Department of Physiology and Cellular Biophysics, Columbia University, New York, New York 10032, USA. <sup>8</sup>Department of Pathology and Cell Biology, Columbia University, New York, New York 10032, USA.





**Figure 1 | Calciurein promotes CRTC2 activation during fasting.** **a**, Effect of Ser/Thr phosphatase inhibitors (okadaic acid (OA), Csa) on CRTC2 dephosphorylation and CRE-Luc reporter activation ( $*P < 0.001$ ;  $n = 3$ ). Gcg, glucagon; Veh, vehicle. **b**, Effect of glucagon on dephosphorylation (left) and activity (right) of wild-type (WT) and calciurein-defective ( $\Delta$ Calna) CRTC2 in hepatocytes ( $*P < 0.001$ ;  $n = 3$ ). **c**, Effect of calciurein A overexpression

(left) or knockdown (right) on CRTC2 dephosphorylation (top), CRE-Luc reporter activity (middle,  $*P < 0.001$ ;  $n = 3$ ), and glucose output (bottom,  $*P < 0.001$ ;  $n = 3$ ) from hepatocytes. US, unspecific. **d**, Effect of hepatic calciurein overexpression (left) or knockdown (right) on CRE-Luc activity, gluconeogenic gene (*Pck1*, *G6pc*) expression, and blood glucose concentrations in 6–8 h fasted mice ( $*P < 0.01$ ;  $n = 5$ ). Data are shown as mean  $\pm$  s.e.m.



**Figure 2 | Glucagon stimulates CRTC2 dephosphorylation via activation of InsP<sub>3</sub>Rs.** **a**, Effect of glucagon (Gcg) on calcium mobilization in hepatocytes by fluorescence imaging. Calcium mobilization and calciurein activation following knockdown of all three InsP<sub>3</sub>R family members shown ( $*P < 0.001$ ;  $n = 3$ ). **b**, Effect of calcium chelator (BAPTA) on CRTC2 dephosphorylation and CRE-Luc activation ( $*P < 0.001$ ;  $n = 3$ ). **c**, Effect of InsP<sub>3</sub>R depletion on CRTC2 dephosphorylation, CRE-Luc activity, and glucose output from hepatocytes ( $*P < 0.001$ ;  $n = 3$ ). **d**, Effect of hepatic InsP<sub>3</sub>R knockdown on CRE-Luc activity, blood glucose, and gluconeogenic gene expression ( $*P < 0.01$ ;  $n = 5$ ). Data are shown as mean  $\pm$  s.e.m.

confirmed the effects of InsP<sub>3</sub>R depletion using hepatocytes from mice with a knockout of the InsP<sub>3</sub>R2 (ref. 10), the predominant InsP<sub>3</sub>R isoform in these cells (Supplementary Fig. 6c–e).

On the basis of these results, InsP<sub>3</sub>Rs would also be expected to modulate fasting glucose production *in vivo*. Decreasing hepatic InsP<sub>3</sub>R expression, either by knockdown of all three InsP<sub>3</sub>Rs in liver or by targeted disruption of the *Insp3r2* gene, reduced fasting CRE-Luc activity, gluconeogenic gene expression, and circulating glucose concentrations, demonstrating the importance of these receptors in glucose homeostasis (Fig. 2d and Supplementary Fig. 7).

We tested whether glucagon modulates InsP<sub>3</sub>R activity through PKA-mediated phosphorylation. Exposure of hepatocytes to glucagon increased the phosphorylation of InsP<sub>3</sub>R1 as well as InsP<sub>3</sub>R2 and InsP<sub>3</sub>R3 by immunoblot assay with phospho-PKA substrate antiserum; these effects were blocked by the PKA inhibitor H89 (Fig. 3a and Supplementary Fig. 8a). Moreover, mutation of serine residues at consensus PKA sites in InsP<sub>3</sub>R1 (Ser 1589, Ser 1756) to alanine completely disrupted InsP<sub>3</sub>R1 phosphorylation in response to glucagon (Fig. 3b). As a result, overexpression of PKA-defective (S1589,1756A) InsP<sub>3</sub>R1 interfered with calcium mobilization and calcineurin activation, and it reduced CRE-Luc activation and glucose secretion from hepatocytes exposed to glucagon (Fig. 3b–d).

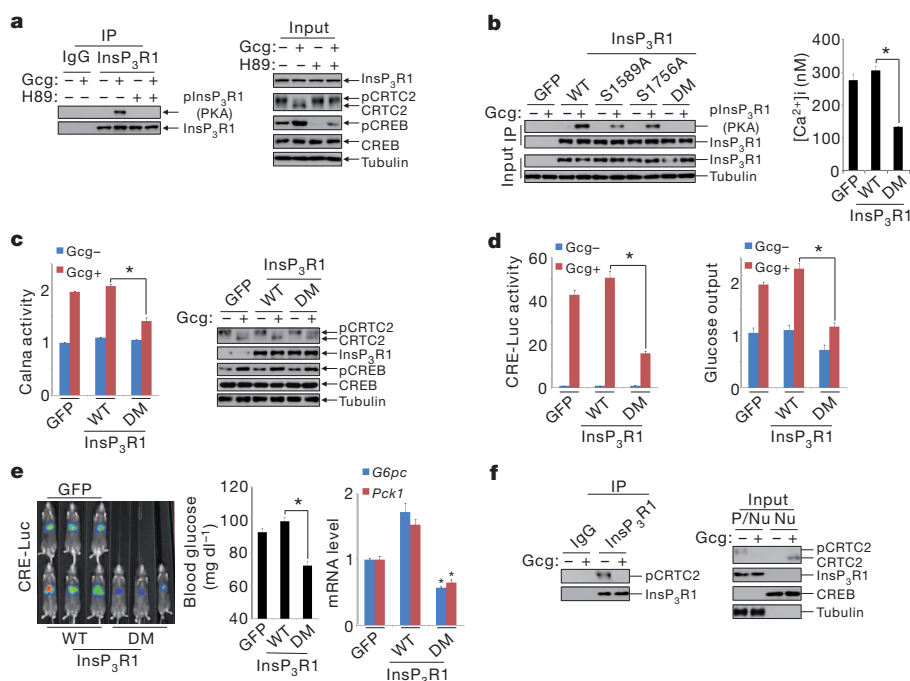
Similar to glucagon, fasting also stimulated hepatic InsP<sub>3</sub>R1 phosphorylation at Ser 1589 and Ser 1756 (Supplementary Fig. 8b). And overexpression of PKA-defective InsP<sub>3</sub>R1 reduced fasting CRE-Luc induction, calcineurin activation, and gluconeogenic gene expression, leading to lower circulating glucose concentrations (Fig. 3e and Supplementary Fig. 8c, d). Taken together, these results support an important role for the PKA-mediated phosphorylation of InsP<sub>3</sub>R in hepatic gluconeogenesis.

We considered that the proximity of CRTC2 to the calcium signaling machinery may be important for its activation. Supporting this

notion, CRTC2 was found to associate with InsP<sub>3</sub>R1 via its amino-terminal CREB binding domain (CBD) in co-immunoprecipitation assays (Fig. 3f and Supplementary Fig. 9a–d). Moreover, CRTC2 was enriched in endoplasmic reticulum containing high density microsomal fractions, which also contain the InsP<sub>3</sub>Rs (Supplementary Fig. 9e). The InsP<sub>3</sub>R–CRTC2 association seems to be critical for CRTC2 localization in the perinuclear space, because RNA interference (RNAi)-mediated knockdown of the InsP<sub>3</sub>Rs led to redistribution of CRTC2 in the cytoplasm (Supplementary Fig. 9f). Disrupting the CRTC2–InsP<sub>3</sub>R interaction, by deletion of the CBD in CRTC2 or by addition of an N-terminal myristoylation signal that targets CRTC2 to the plasma membrane, blocked CRTC2 dephosphorylation and CRE-reporter activation in response to glucagon (Supplementary Fig. 9g–i). Taken together, these results suggest that the association of CRTC2 with InsP<sub>3</sub>Rs enhances its sensitivity to fasting signals.

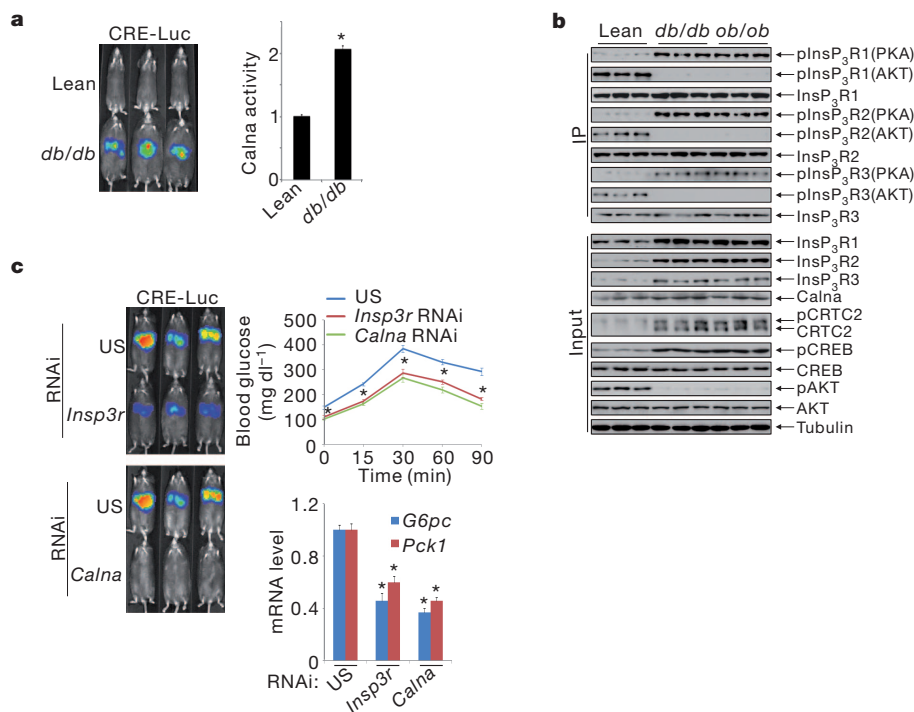
Under feeding conditions, insulin inhibits gluconeogenesis in part by increasing CRTC2 phosphorylation. We wondered whether insulin interferes with InsP<sub>3</sub>R effects on CRTC2 activation. Supporting this idea, AKT has been shown to block calcium mobilization by phosphorylating InsP<sub>3</sub>Rs at Ser 2682 (in InsP<sub>3</sub>R1)<sup>11</sup>. Indeed, exposure of hepatocytes to insulin increased InsP<sub>3</sub>R phosphorylation by immunoblot analysis with phospho-AKT substrate antiserum (Supplementary Fig. 10a); mutation of Ser 2682 (in InsP<sub>3</sub>R1) to alanine blocked these effects. Insulin treatment also reduced glucagon-dependent increases in calcium mobilization and calcineurin activation in cells expressing wild-type InsP<sub>3</sub>R1, but it had no effect in cells expressing AKT-defective (S2682A) InsP<sub>3</sub>R1 (Supplementary Fig. 10b). As a result, CRTC2 dephosphorylation, CRE-Luc activity, and glucose output were elevated in hepatocytes expressing InsP<sub>3</sub>R(S2682A) compared to wild type (Supplementary Fig. 10c).

We examined whether InsP<sub>3</sub>R1 phosphorylation by AKT is important in regulating hepatic glucose production *in vivo*. In line with this



**Figure 3 | Glucagon stimulates CRTC2 activity via PKA-dependent phosphorylation of InsP<sub>3</sub>Rs.** **a**, **b**, Immunoblots of InsP<sub>3</sub>R1 immunoprecipitates (IP) using phospho-PKA substrate antiserum to show effect of H89 (**a**) and Ala mutations (**b**) at one or both (DM) PKA consensus sites (Ser 1589, Ser 1756) on InsP<sub>3</sub>R1 phosphorylation in hepatocytes exposed to glucagon (Gcg). Effect of wild-type and PKA-mutant InsP<sub>3</sub>R1 on calcium mobilization in response to Gcg (**b**) shown (\**P* < 0.001; *n* = 3). **c**, **d**, Effect of wild-type or PKA-defective InsP<sub>3</sub>R1 (DM) on calcineurin (Calna) activation

(**c**) and CRTC2 dephosphorylation (**c**), as well as CRE-Luc activation (**d**) and glucose output (**d**) from hepatocytes (\**P* < 0.001; *n* = 3). **e**, Effect of wild-type and PKA-defective InsP<sub>3</sub>R1 on hepatic CRE-Luc activity, fasting blood glucose, and gluconeogenic gene expression (*G6pc*, *Pck1*) (\**P* < 0.01 versus wild type; *n* = 5). **f**, Co-immunoprecipitation of CRTC2 with InsP<sub>3</sub>R1 in primary hepatocytes. Exposure to glucagon (100 nM, 15 min) indicated. Input levels of CRTC2 and InsP<sub>3</sub>R1 in nuclear (Nu) and post-nuclear (p/Nu) supernatant fractions shown. Data are shown as mean ± s.e.m.



**Figure 4 | InsP<sub>3</sub>R activity is upregulated in diabetes.** **a**, Hepatic CRE-Luc and calcineurin activity in lean and *db/db* mice (\*P < 0.001; n = 5). **b**, Immunoblots showing relative amounts and phosphorylation of InsP<sub>3</sub>R family members in livers of ad libitum fed lean, *db/db*, or *ob/ob* mice. InsP<sub>3</sub>R

phosphorylation at PKA or AKT sites indicated. **c**, Effect of RNAi-mediated depletion of InsP<sub>3</sub>Rs or calcineurin A on CRE-Luc activity, gluconeogenic gene expression, and hepatic glucose production in *db/db* mice, determined by pyruvate tolerance testing (\*P < 0.01; n = 5). Data are shown as mean ± s.e.m.

notion, feeding increased hepatic InsP<sub>3</sub>R1 phosphorylation at Ser 2682 (Supplementary Fig. 8b). Moreover, overexpression of AKT-defective InsP<sub>3</sub>R1 partially suppressed feeding-induced decreases in CRE-Luc activity and gluconeogenic gene expression, leading to elevations in circulating glucose concentrations (Supplementary Fig. 10d). Taken together, these results suggest that the AKT-mediated phosphorylation of InsP<sub>3</sub>Rs during feeding inhibits hepatic gluconeogenesis by blocking the calcineurin-dependent dephosphorylation of CRTC2.

We wondered whether hepatic InsP<sub>3</sub>R signalling contributes to increases in gluconeogenesis in the setting of insulin resistance. Supporting this notion, hepatic calcineurin activity was enhanced in both *ob/ob* and *db/db* diabetic animals, leading to increases in CRE-Luc activity (Fig. 4a and Supplementary Fig. 11a, b). Pointing to a role for InsP<sub>3</sub>R, hepatic amounts of PKA-phosphorylated, active InsP<sub>3</sub>Rs were increased in these diabetic mice, whereas amounts of AKT-phosphorylated, inactive InsP<sub>3</sub>Rs were reduced (Fig. 4b). Correspondingly, knockdown of either calcineurin or InsP<sub>3</sub>Rs in *db/db* mice reduced CRE-Luc activity, gluconeogenic gene expression, and hepatic gluconeogenesis (Fig. 4c and Supplementary Fig. 11c).

Collectively, our results demonstrate that glucagon promotes CRTC2 dephosphorylation during fasting by triggering increases in cytoplasmic calcium that lead to calcineurin activation (Supplementary Fig. 12). The ability for glucagon to increase calcium signalling via the PKA-mediated phosphorylation of InsP<sub>3</sub>Rs demonstrates an important regulatory node for cross-talk between cAMP and calcium signalling pathways in liver and perhaps other insulin sensitive tissues. The partial inhibition of calcium entry by the PKA inhibitor H89 also points to additional regulatory inputs<sup>12,13</sup> that may function with PKA to increase InsP<sub>3</sub>R activity in response to glucagon. CRTC2 has also been found to stimulate metabolic gene expression by upregulating the nuclear hormone receptor co-activator PGC-1 $\alpha$  in liver<sup>14,15</sup> and muscle<sup>16</sup>. On the basis of the well-recognized role of calcium signalling in PGC-1 $\alpha$ -dependent transcription, InsP<sub>3</sub>Rs may also have an important function in this setting.

## METHODS SUMMARY

Adenoviruses were delivered by tail vein injection<sup>17</sup>. Hepatic CRE-Luc activity was visualized using an IVIS Imaging system. Mice were imaged 3–5 days after injection of CRE-Luc adenovirus. Pyruvate tolerance testing was performed on mice fasted overnight and injected intraperitoneally with pyruvate (2 g kg<sup>-1</sup>). *Insp3r2* knockout mice have been described<sup>10</sup>. Cultured primary mouse hepatocytes were prepared as reported<sup>18</sup>. Cellular fractionation studies were conducted using primary mouse hepatocytes<sup>18</sup>. Calcium imaging experiments were performed using a CCD camera on primary hepatocytes loaded with fura-2 dye. Mass spectrometry studies were performed on CRTC2 immunoprecipitates prepared from HEK293T cells and on immunoprecipitates of phospho-PKA substrate antiserum prepared from primary hepatocytes exposed to glucagon. Anti-InsP<sub>3</sub>R1 (A302–158A) and InsP<sub>3</sub>R3 (A302–160A) antibodies were purchased from Bethyl Laboratories, anti-InsP<sub>3</sub>R2 (ab77838) antiserum was from Abcam, anti-calcineurin (610260) from BD Biosciences, anti-GRP78 (ADI-SPA-826-F) from Enzo Life Sciences, anti-phospho-PKA substrate (RRXS/T, 9624), anti-phospho-AKT substrate (RRXS/T, 9614) and CRTC2 (pS171, 2892) from Cell Signaling. Phospho (Ser 275) CRTC2 antibody was used as described<sup>19</sup>. For more details, see Supplementary Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 26 July 2011; accepted 22 February 2012.

Published online 8 April 2012.

- Altarejos, J. Y. & Montminy, M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. *Nature Rev. Mol. Cell Biol.* **12**, 141–151 (2011).
- Yoon, Y. S. *et al.* Suppressor of MEK null (SMEK)/protein phosphatase 4 catalytic subunit (PP4C) is a key regulator of hepatic gluconeogenesis. *Proc. Natl Acad. Sci. USA* **107**, 17704–17709 (2010).
- Screaton, R. A. *et al.* The CREB coactivator TORC2 functions as a calcium- and cAMP-sensitive coincidence detector. *Cell* **119**, 61–74 (2004).
- Hogan, P. G., Chen, L., Nardone, J. & Rao, A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev.* **17**, 2205–2232 (2003).
- Ferris, C. D., Haganir, R. L., Bredt, D. S., Cameron, A. M. & Snyder, S. H. Inositol trisphosphate receptor: phosphorylation by protein kinase C and calcium calmodulin-dependent protein kinases in reconstituted lipid vesicles. *Proc. Natl Acad. Sci. USA* **88**, 2232–2235 (1991).
- Volpe, P. & Alderson-Lang, B. H. Regulation of inositol 1,4,5-trisphosphate-induced Ca<sup>2+</sup> release. II. Effect of cAMP-dependent protein kinase. *Am. J. Physiol.* **258**, C1086–C1091 (1990).



7. Bird, G. S., Burgess, G. M. & Putney, J. W. Jr. Sulfhydryl reagents and cAMP-dependent kinase increase the sensitivity of the inositol 1,4,5-trisphosphate receptor in hepatocytes. *J. Biol. Chem.* **268**, 17917–17923 (1993).
8. Patterson, R. L., Boehning, D. & Snyder, S. H. Inositol 1,4,5-trisphosphate receptors as signal integrators. *Annu. Rev. Biochem.* **73**, 437–465 (2004).
9. Futatsugi, A. *et al.* IP3 receptor types 2 and 3 mediate exocrine secretion underlying energy metabolism. *Science* **309**, 2232–2234 (2005).
10. Cruz, L. N. *et al.* Regulation of multidrug resistance-associated protein 2 by calcium signaling in mouse liver. *Hepatology* **52**, 327–337 (2010).
11. Szado, T. *et al.* Phosphorylation of inositol 1,4,5-trisphosphate receptors by protein kinase B/AKT inhibits  $\text{Ca}^{2+}$  release and apoptosis. *Proc. Natl Acad. Sci. USA* **105**, 2427–2432 (2008).
12. Tovey, S. C. *et al.* Regulation of inositol 1,4,5-trisphosphate receptors by cAMP independent of cAMP-dependent protein kinase. *J. Biol. Chem.* **285**, 12979–12989 (2010).
13. Wakelam, M. J., Murphy, G. J., Hruby, V. J. & Houslay, M. D. Activation of two signal-transduction systems in hepatocytes by glucagon. *Nature* **323**, 68–71 (1986).
14. Yoon, J. *et al.* Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001).
15. Herzig, S. *et al.* CREB regulates hepatic gluconeogenesis via the co-activator PGC-1. *Nature* **413**, 179–183 (2001).
16. Wu, Z. *et al.* Transducer of regulated CREB-binding proteins (TORCs) induce PGC-1 $\alpha$  transcription and mitochondrial biogenesis in muscle cells. *Proc. Natl Acad. Sci. USA* **103**, 14379–14384 (2006).
17. Dentin, R. *et al.* Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2. *Nature* **449**, 366–369 (2007).
18. Wang, Y., Vera, L., Fischer, W. H. & Montminy, M. The CREB coactivator CRTC2 links hepatic ER stress and fasting gluconeogenesis. *Nature* **460**, 534–537 (2009).
19. Jansson, D. *et al.* Glucose controls CREB activity in islet cells via regulated phosphorylation of TORC2. *Proc. Natl Acad. Sci. USA* **105**, 10161–10166 (2008).
20. Liu, Y. *et al.* A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange. *Nature* **456**, 269–273 (2008).
21. Wang, B. *et al.* A hormone-dependent module regulating energy balance. *Cell* **145**, 596–606 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by National Institutes of Health grants R01-DK049777, R01-DK083834 and R01-DK091618 (M.M.), HL087123 (I.T.), the Kieckhefer Foundation, The Clayton Foundation for Medical Research, and the Leona M. and Harry B. Helmsley Charitable Trust.

**Author Contributions** Y.W., I.T. and M.M. designed and interpreted the experiments. Y.W., G.L., J.C.P., R.S. and J.G. carried out the experimental work. Y.W., K.O. and J.C. characterized glucose metabolism in *Insp3r2* knockout mice. W.H.F. performed proteomic studies, and Y.W. and M.M. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.M. ([montminy@salk.edu](mailto:montminy@salk.edu)).

## METHODS

**Mouse strains and adenovirus.** Adenoviruses ( $1 \times 10^8$  plaque forming units (p.f.u.) GFP, calcineurin, *InsP<sub>3</sub>R1*, *InsP<sub>3</sub>R1* DM (S1589A/S1756A), unspecific RNAi, calcineurin RNAi, *Insp3r1* RNAi, *Insp3r2* RNAi, *Insp3r3* RNAi, *Crtc2* RNAi,  $1 \times 10^9$  p.f.u. CRE-Luc reporter,  $5 \times 10^7$  p.f.u. RSV  $\beta$ -galactosidase) were delivered to 8–10-week-old male C57BL/6J, B6.V-*lep* < ob>/J, B6.Cg-*m*+/*+* *Lepr* < db>/J mice by tail vein injection<sup>17</sup>. *Insp3r2* knockout mice were described previously<sup>10</sup>. All mice were adapted to their environment for 1 week before study and were housed in colony cages with a 12 h light/dark cycle in a temperature-controlled environment. For *in vivo* imaging experiments, mice were imaged on day 3–5 after adenovirus delivery. Wild-type CRTC2, CRTC2(S171A), GFP, unspecific RNAi, *Crtc2* RNAi, CRE-Luc and RSV  $\beta$ -gal adenoviruses have been described previously<sup>17,20</sup>. The adenoviruses containing rat *InsP<sub>3</sub>R1*, *InsP<sub>3</sub>R1* DM and *InsP<sub>3</sub>R1*(S2682A) were generated from the *InsP<sub>3</sub>R1* plasmid, provided by I. Bezprozvanny (UT Southwestern Medical Center at Dallas). Calcineurin adenovirus was constructed using a mouse calcineurin plasmid (Addgene). CRTC2  $\Delta$ CBD (51–692 amino acids), S275A and S171A/S275A adenoviruses were made from mouse CRTC2. Myristoylated CRTC2 (Myr-CRTC2) adenovirus was generated with mouse CRTC2 fused to an N-terminal myristoylation tag (MGSSSKPKDPSQR) from Src. Calcineurin RNAi, *Insp3r1* RNAi, *Insp3r2* RNAi, *Insp3r3* RNAi adenoviruses were constructed using the sequence 5'-GGGTACCGCATGTACAGGAAAA-3', 5'-GGGTACTGGAATAGCCTCTTCC-3', 5'-GGGTAACAAGCACCACCATCCC-3' and 5'-GGGCAAGCTGCA GGTGTTCTTG-3', respectively. All expressed constructs used in this study were confirmed by sequencing.

**In vivo analysis.** For *in vivo* imaging, mice were imaged as described<sup>17,20</sup> under ad libitum feeding conditions or after fasting for 6 h. For pyruvate challenge experiments, mice were fasted overnight and injected intraperitoneally with pyruvate ( $2 \text{ g kg}^{-1}$ ). Blood glucose values were determined using a LifeScan automatic glucometer. For immunoblot, mouse tissues were sonicated, centrifuged and supernatants were reserved for protein determinations, and SDS-PAGE analysis.

**In vitro analysis.** HEK293T (ATCC) cells were cultured in DMEM containing 10% FBS (HyClone),  $100 \text{ mg ml}^{-1}$  penicillin-streptomycin. Mouse primary hepatocytes were isolated and cultured as previously described<sup>18</sup>. Cellular fractionation studies were conducted as previously reported<sup>18</sup>. For reporter studies, Ad-CRE-Luc-infected hepatocytes ( $1 \text{ p.f.u. per cell}$ ) were exposed to glucagon ( $100 \text{ nM}$ ) for 2 to  $\sim 4 \text{ h}$ . For CsA ( $10 \text{ }\mu\text{M}$ ), okadaic acid ( $100 \text{ nM}$ ), cell permeable calcineurin autoinhibitory peptide ( $10 \text{ }\mu\text{M}$ ), CN585 ( $100 \text{ }\mu\text{M}$ ), calyculin A ( $10 \text{ nM}$ ), xestospongin C ( $2 \text{ }\mu\text{M}$ ), H89 ( $30 \text{ }\mu\text{M}$ ) or BAPTA ( $50 \text{ }\mu\text{M}$ ) inhibition, hepatocytes were pre-treated with the inhibitors for 1 h. Luciferase activities were normalized to  $\beta$ -galactosidase activity from adenoviral-encoded RSV  $\beta$ -galactosidase.

Calcineurin activity (test kit from Enzo Life Sciences) and cellular cAMP levels (test kit from Cayman Chemical Company) were measured according to the manufacturer's instructions.

**Calcium imaging.** Mouse primary hepatocytes were plated on glass coverslips and loaded with  $5 \text{ }\mu\text{M}$  Fura-2 acetoxymethyl ester (Molecular Probes) in the presence of 0.025% (w/v) pluronic F127 (Sigma-Aldrich) in Media 199 (Mediatech) for 30 min. Coverslips were mounted on a laminar flow perfusion chamber (Warner Instruments) and perfused with Media 199 or a solution of  $100 \text{ nM}$  glucagon in Media 199. Images of Fura-2 loaded cells were collected with a cooled CCD camera while the excitation wavelength was alternated between 340 nm and 380 nm. The ratio of fluorescence intensity at the two excitation wavelengths was calculated after subtracting background fluorescence.  $[\text{Ca}^{2+}]_i$  (cytosolic free calcium concentration) was calculated using a Fura-2 calcium imaging calibration kit (Invitrogen). Images were collected and analysed using the MetaFluor software package (Universal Imaging Corp.). Graphs represent average responses from groups of 30–40 individual cells from representative single experiments. Bar graphs represent average responses (fold over average baseline) from 150–200 cells per condition. All experiments were repeated at least three times with similar results.

**Immunoblot, immunoprecipitation and immunostaining.** Immunoblot, immunoprecipitation and immunostaining assays were performed as described<sup>18</sup>. CRTC2, pCREB (Ser 133), CREB, pAKT (Thr 308), AKT, tubulin, HA and Flag antibodies were previously described<sup>18</sup>. The antibodies anti-*InsP<sub>3</sub>R1* (A302–158A) and *InsP<sub>3</sub>R3* (A302–160A) were purchased from Bethyl Laboratories, anti-*InsP<sub>3</sub>R2* (ab77838) from Abcam, anti-calcineurin (610260) from BD Biosciences, anti-GRP78 (ADI-SPA-826-F) from Enzo Life Sciences, anti-phospho-PKA substrate (RRXS/T, 9624), anti-phospho-AKT substrate (RXXS/T, 9614) and CRTC2 (pS171, 2892) from Cell Signaling. CRTC2 (pS275) antibody was used as described<sup>19</sup>.

**Quantitative PCR.** Total cellular RNAs from whole liver or from primary hepatocytes were extracted using the RNeasy kit (Qiagen) and used to generate cDNA with SuperScript II enzyme (Invitrogen). cDNA were analysed by quantitative PCR as described<sup>18</sup>.

**Mass spectrometry.** Immunoprecipitates of endogenous CRTC2 from HEK293T cells and of phospho-PKA substrate antiserum from glucagon-stimulated hepatocytes were prepared for mass spectrometric studies as previously reported<sup>21</sup>, and analysed by electrospray ionization tandem mass spectrometry on a Thermo LTQ Orbitrap instrument.

**Statistical analyses.** All studies were performed on at least three independent occasions. Results are reported as mean  $\pm$  s.e.m. The comparison of different groups was carried out using two-tailed unpaired Student's *t*-test. Differences were considered statistically significant at  $P < 0.05$ .

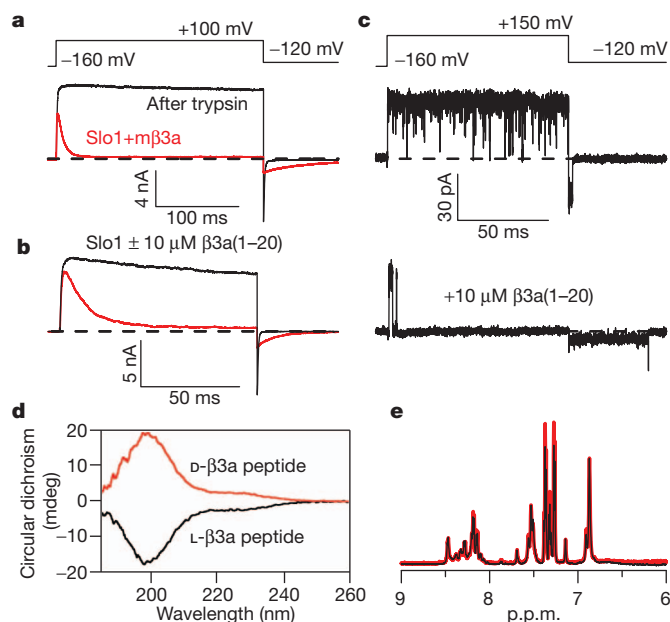
# Stereospecific binding of a disordered peptide segment mediates BK channel inactivation

Vivian Gonzalez-Perez<sup>1\*</sup>, Xu-Hui Zeng<sup>1\*†</sup>, Katie Henzler-Wildman<sup>2</sup> & Christopher J. Lingle<sup>1</sup>

A number of functionally important actions of proteins are mediated by short, intrinsically disordered peptide segments<sup>1</sup>, but the molecular interactions that allow disordered domains to mediate their effects remain a topic of active investigation<sup>2–5</sup>. Many K<sup>+</sup> channel proteins, after initial channel opening, show a time-dependent reduction in current flux, termed ‘inactivation’, which involves movement of mobile cytosolic peptide segments (approximately 20–30 residues) into a position that physically occludes ion permeation<sup>6–8</sup>. Peptide segments that produce inactivation show little amino-acid identity<sup>6,9–13</sup> and tolerate appreciable mutational substitutions<sup>13</sup> without disrupting the inactivation process. Solution nuclear magnetic resonance of several isolated inactivation domains reveals substantial conformational heterogeneity with only minimal tendency to ordered structures<sup>14–17</sup>. Channel inactivation mechanisms may therefore help us to decipher how intrinsically disordered regions mediate functional effects. Whereas many aspects of inactivation of voltage-dependent K<sup>+</sup> channels (Kv) can be described by a simple one-step occlusion mechanism<sup>6,7,18,19</sup>, inactivation of the voltage-dependent large-conductance Ca<sup>2+</sup>-gated K<sup>+</sup> (BK) channel mediated by peptide segments of auxiliary  $\beta$ -subunits involves two distinguishable kinetic steps<sup>20,21</sup>. Here we show that two-step inactivation mediated by an intrinsically disordered BK  $\beta$ -subunit peptide involves a stereospecific binding interaction that precedes blockade. In contrast, blocking mediated by a *Shaker* Kv inactivation peptide is consistent with direct, simple occlusion by a hydrophobic segment without substantial steric requirement. The results indicate that two distinct types of molecular interaction between disordered peptide segments and their binding sites produce qualitatively similar functions.

For BK channels, inactivation is mediated by cytosolic amino-terminal segments of auxiliary  $\beta$ -subunits (Supplementary Fig. 1)<sup>20–24</sup>. In contrast to simple models proposed for Kv N-type inactivation ( $C \rightleftharpoons O \rightleftharpoons I$ ) for BK  $\beta$ -subunits the mechanism of inactivation involves two steps<sup>20,21</sup> ( $C \rightleftharpoons O^* \rightleftharpoons I^*$ ), an initial transition to a pre-inactivated conducting state ( $O^*$ ) followed by transition to a non-conducting inactivated state ( $I^*$ ) (Supplementary Fig. 2a). The properties of inactivation onset and recovery differ among subunits<sup>20–24</sup>, presumably dependent on the interactions of each  $\beta$ -subunit N terminus and the BK  $\alpha$ -subunit. The site(s) and physical basis for this sequence of steps in BK inactivation are not understood. Here we exploit inactivation mediated by the BK  $\beta$ 3a subunit to tease apart the molecular steps in two-step inactivation<sup>21</sup>. First, both the  $\beta$ 3a subunit (Fig. 1a) and a  $\beta$ 3a peptide (Fig. 1b) produce a slowing of BK tail currents that arises from buffering of channels between  $O^* \rightleftharpoons I^*$ . Net current flux during the tail current greatly exceeds that expected for a population of open BK channels simply passing back to closed states<sup>21</sup> (Fig. 1a, b). Second, the unique properties of the  $O^* \rightleftharpoons I^*$  equilibrium can be directly visualized in recordings of single BK channels during recovery from inactivation mediated either by  $\beta$ 3a N terminus<sup>21</sup> or by  $\beta$ 3a peptides

(Fig. 1c). After repolarization, inactivated single channels open to a current level of reduced amplitude, higher variance and longer duration than a typical BK channel closing (Fig. 1c). The lower current level is a direct reflection of the rapid voltage-dependent equilibrium between states  $O^*$  and  $I^*$  (Supplementary Fig. 2). Because the isolated  $\beta$ 3a peptide mimics the effects of the tethered  $\beta$ 3a N terminus (Fig. 1a–c), the peptide can be exploited for examination of the structural and functional basis of the two-step mechanism. Circular dichroism spectra of both a  $\beta$ 3a peptide composed of the naturally occurring L-amino-acids and an unnatural, mirror-image peptide composed of D-amino-acid enantiomers (Fig. 1d) are typical of random coil and an absence of secondary structure. Solution nuclear magnetic resonance (NMR) spectra of these two peptides are identical and show minimal chemical shift dispersion (Fig. 1e and Supplementary Fig. 3a). The peptides also show very few inter-residue nuclear Overhauser effects (NOEs) in



**Figure 1 | Intrinsically disordered  $\beta$ 3a N-terminal peptide mimics the two-step inactivation produced by the tethered  $\beta$ 3a N terminus.** **a**, Intact  $\beta$ 3a subunits mediate trypsin-sensitive inactivation and tail current prolongation of BK channels. Red trace, BK  $\alpha$  +  $\beta$ 3a currents with the indicated protocol and 10  $\mu$ M Ca<sup>2+</sup>; black trace, after trypsin. **b**,  $\beta$ 3a(1–20) peptide (red) produces inactivation and slows BK tail currents. **c**, Single BK channel in the absence (upper) and presence (lower) of  $\beta$ 3a(1–20) peptide. Peptide produces block and a noisy, long-duration low-amplitude opening before unblock. **d**, Circular dichroism spectra of L- (black) and D- $\beta$ 3a-peptides (red) are identical except for expected opposite sign. Minimum below 200 nm is characteristic of random coil conformation. **e**, Solution NMR of  $\beta$ 3a 21-mer peptides (black, L-peptide; red, D-peptide) shows minimal chemical shift dispersion in the amide region.

<sup>1</sup>Department of Anesthesiology, Washington University School of Medicine, St Louis, Missouri 63110, USA. <sup>2</sup>Department of Biochemistry, Washington University School of Medicine, St Louis, Missouri 63110, USA. <sup>†</sup>Present address: Institute of Life Science, Nanchang University, 999 Xuefu Road, Honggu District, Nanchang 330031, China.

\*These authors contributed equally to this work.



two-dimensional homonuclear nuclear Overhauser effect spectroscopy (NOESY) spectra (Supplementary Fig. 3b), consistent with intrinsic disorder.

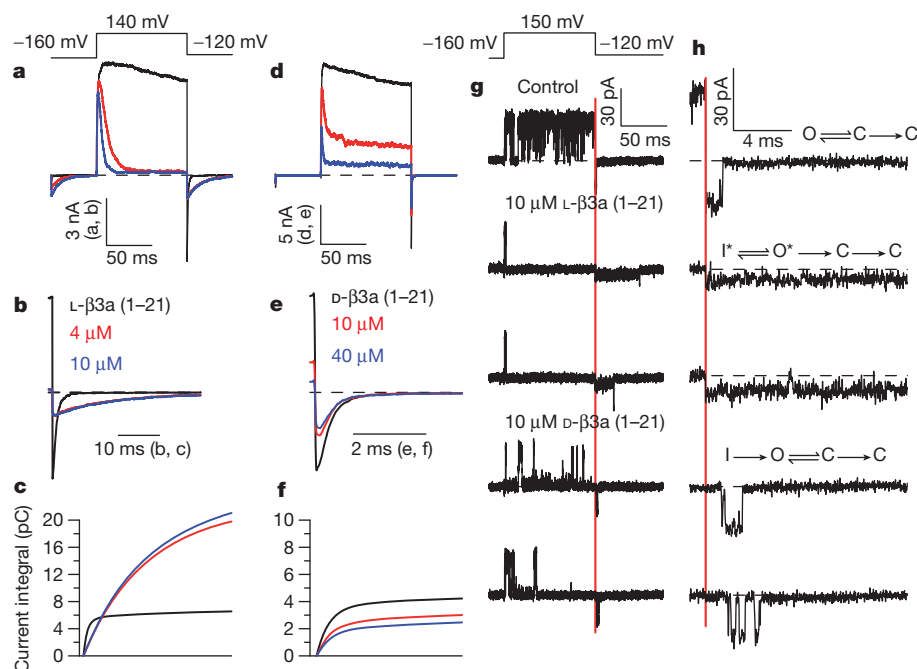
We propose that the two-step process involves a binding step that precedes rapid motions of a flexible part of the bound N terminus in and out of a position of pore occlusion. If binding has a unique steric requirement (stereospecific binding), a  $\beta$ 3a peptide composed entirely of D-amino acids should fail to interact with this site. Both 21-mer L- and D-amino-acid  $\beta$ 3a peptides produced block of BK current (Fig. 2a, d), with block by the D-peptide of lower affinity (Supplementary Fig. 4). However, the D-peptide (Fig. 2e, f) failed to mimic the ability of the L-peptide (Fig. 2b, c) to slow the tail current decay and increase tail current flux.

Because the D-peptide acts with lower affinity, the differential effects of the two peptides might arise, not from a difference in underlying mechanism, but because the O\* state lifetime involving the D-peptide may be too brief to resolve. We therefore compared the ability of L- and D-peptides to block single-channel BK openings (Fig. 2g). With 10  $\mu$ M D-peptide, we observed frequent unblocking events, consistent with its weaker affinity. After repolarization, a channel blocked by the L-peptide shows rapid unblock to the reduced current level, before unblock to a full BK open level. In contrast, a channel blocked by the D-peptide re-opens to a full current level, with no hint of the reduced current level seen with the L-peptide. When examined at higher time resolution (Fig. 2h), repolarization after L-peptide block resulted in an essentially instantaneous step to the level of reduced current, consistent with the properties of the O\*-I\* equilibrium (Supplementary Fig. 2d-f). Repolarization after D-peptide block revealed a finite delay before a full BK opening (Fig. 2h), similar to block of Kv channels<sup>18</sup>. The differential effects of the peptides show that stereospecific binding of L-peptide is required to produce the tail current prolongation. Although the steric binding occurs even in the absence of intrinsic structure in the free peptide, it would not be surprising that, once bound, the peptide adopts a specific structure. We also found that inactivation produced by a  $\beta$ 3a N terminus

( $\beta$ 3a-R16Q/R17Q/R18Q) in which the first 20 residues are uncharged (Supplementary Fig. 5) retains the slow tail currents and single-channel openings indicative of two-step inactivation. We conclude that the binding event underlying the O\* state depends on hydrophobic interactions with unique steric requirements without significant contribution from electrostatic effects.

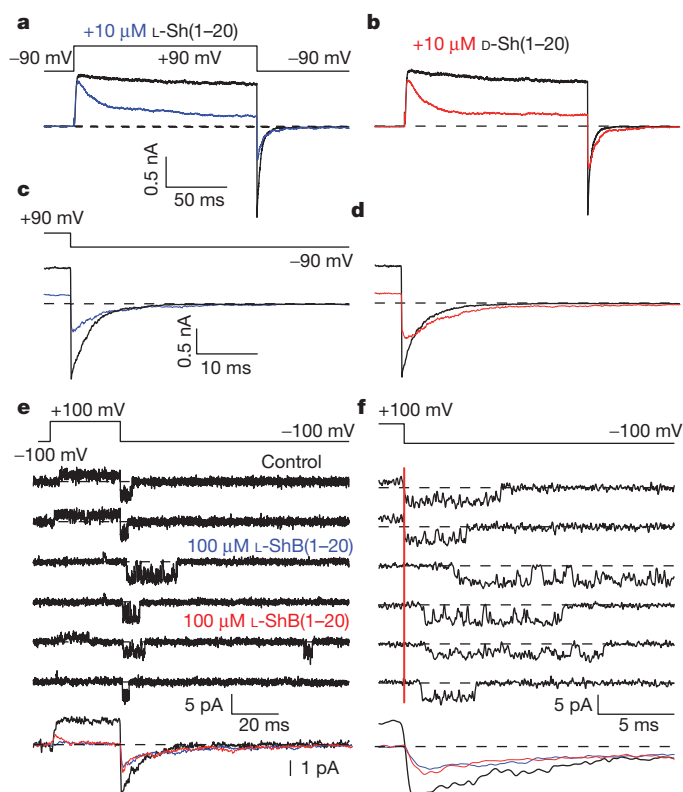
If the D-peptide blocking events arise from relatively non-specific binding, not requiring a specific steric fit but dependent on bulk hydrophobicity, the L-peptide should occasionally act in a similar fashion, in accordance with relative occupancies of non-specific and stereospecific sites. For a set of single BK channel tail openings in the presence of 10  $\mu$ M  $\beta$ 3a L-peptide, approximately 3% of tail current re-openings occurred with a well-defined latency before opening to a full open channel level; openings to a reduced current level with no latency were approximately 30-fold more abundant (Supplementary Fig. 6a, b). Thus, the BK channel can be blocked in two distinct ways by the  $\beta$ 3a L-peptide. The more common re-opening to the reduced current level (O\*-I\*) is only observed with the L-peptide. However, re-opening to a fully open level is observed more rarely with L-peptide, and this behaviour is similar to the one-step block (O-I) characteristic of the  $\beta$ 3a D-peptide or block of the *Shaker* channel<sup>18</sup>. Because recovery from I\* is much more rapid (Supplementary Fig. 2) than for the rarer form of block (O-I), the peptide position within the channel must differ markedly in each case.

We next asked whether peptide blockade of *Shaker* Kv channels may show a similar steric dependence. *Shaker* inactivation peptides, which also show little defined structure<sup>14</sup>, tolerate appreciable amino-acid substitution without altering inactivation as long as bulk hydrophobicity is maintained<sup>13</sup>. If block largely reflects bulk hydrophobicity with no steric requirement, L- and D-*Shaker* peptides should show similar block affinities and kinetics. Tests of 20-mer L- and D-*Shaker* peptides on *Shaker*( $\Delta$ 2-46) channels confirmed this expectation with little difference in blocking affinity (Fig. 3a-d and Supplementary Fig. 7). Unblock from either peptide occurred in a fashion consistent with a single-step occlusion mechanism not dependent on stereochemistry



**Figure 2 | Intrinsically disordered D- and L- $\beta$ 3a-peptides block BK channels, but only the L-peptide produces unique tail current behaviour.** **a**, BK current was activated with 10  $\mu$ M cytosolic  $\text{Ca}^{2+}$  with the indicated voltage protocol; 0 (black), 4 (red) and 10 (blue)  $\mu$ M L- $\beta$ 3a(1-21) peptide. **b**, Effects of 4 and 10  $\mu$ M L-peptide shown on a faster time base. **c**, Current integrals of tail currents from panel **b**. Same time base as in **b**. **d**, Currents from another patch with 0, 10 and

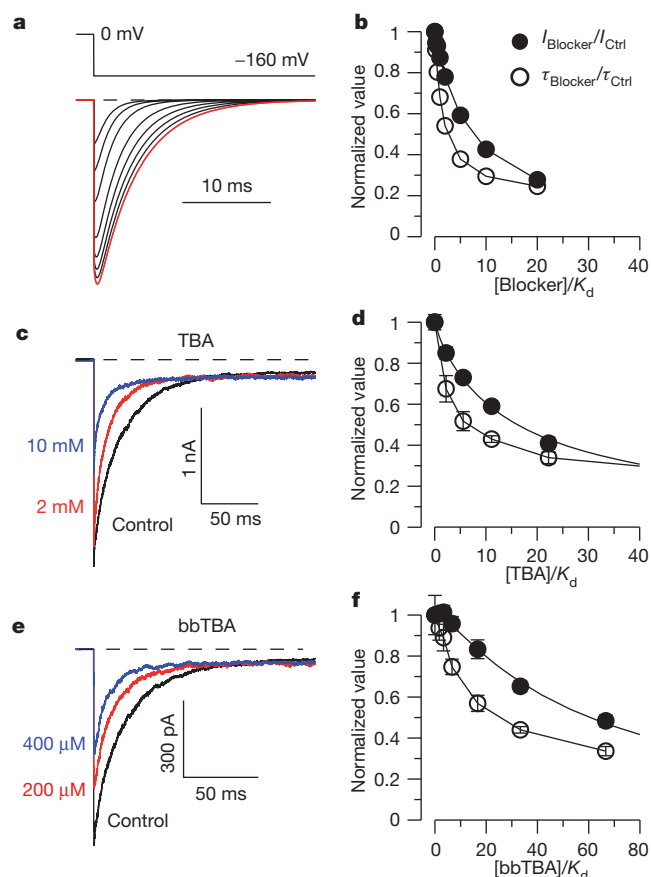
40  $\mu$ M D- $\beta$ 3a(1-21)-peptide. **e**, Tail currents from **d** on a faster time base. **f**, Current integrals of tail currents from **e**. Same time base as **e**. **g**, Single BK channel (10  $\mu$ M  $\text{Ca}^{2+}$ ) showing control trace, two traces with 10  $\mu$ M L-peptide and two traces with 10  $\mu$ M D-peptide. **h**, Faster time base examples of traces in **g** highlighting differences in tail current openings, consistent with indicated models. Red bar, time of repolarization.



**Figure 3** | L- and D-Shaker peptides block Shaker-IR channels in a similar fashion. **a**, Currents in inside-patches activated with the indicated voltage protocol with (blue) and without (black) 10  $\mu$ M L-Shaker peptide. **b**, Currents with (red) and without (black) 10  $\mu$ M D-Shaker peptide. **c**, Tail currents from **a** at expanded time base. **d**, Tail currents from **b** at expanded time base; **a** and **b** are from the same patch. **e**, A single Shaker inactivation-removed (IR) channel was activated and tail currents monitored during application of 100  $\mu$ M L- or D-Shaker peptides. Bottom panel shows averages of 95 sweeps for no peptide (black), and 105 sweeps for D- (red) and L-peptides (blue). **f**, Traces from **e** at higher time resolution. Red bar, time of repolarization. Tail openings after block by D- or L-peptide occur with an average delay of 0.51 ms (97 openings) and 0.57 ms (95 openings), respectively.

(Fig. 3e, f). Thus, the key determinants that define block behaviour of the Shaker ball peptide within the Shaker inner pore do not include a steric interaction. However, this does not exclude that a Shaker peptide might electrostatically interact with positions that may influence block affinity<sup>13,25</sup> or that a form of two-step inactivation also proposed in Kv channels<sup>8,25</sup> may occur (see Supplementary Text).

Our results do not identify the location of the  $\beta$ 3a L-peptide binding site. For Shaker K<sup>+</sup> channels, the ability of TEA to slow inactivation onset argues that TEA and the Shaker N terminus compete for a common position within the inner pore<sup>26</sup>. For BK inactivation, the inability of pore blockers to interfere with inactivation onset is a natural consequence of the two-step inactivation mechanism<sup>20</sup>. If a fast blocker binds with similar affinity to both O and O\* states and the rate-limiting inactivation step is the transition from O to O\*, even if a blocker competitively inhibits movement of an inactivation domain into a position of inactivation, it will not slow the onset of inactivation (Supplementary Fig. 8). However, the effects of blockers on slow tail current allow mechanistically informative tests of the interactions of pore blockers and the inactivation domain (Supplementary Fig. 8). We considered three cases: one in which blocker competitively impedes formation of O\* and subsequently I\* (Supplementary Fig. 8a, model 1), one in which blocker does not impede any aspect of the inactivation process (Supplementary Fig. 8a, model 2) and two variants of a model (Supplementary Fig. 8a, models 3 and 3') in which block competitively inhibits inactivation. Each scheme results in distinct predictions about



**Figure 4** | BK pore blockers compete with inactivation, but not N terminus binding. **a**, Tail currents simulated from model 3 (Supplementary Fig. 8a, model 3) as blocker concentration increases. **b**, Model predictions for normalized tail current amplitude ( $I_{\text{Blocker}}/I_{\text{Ctrl}}$ ) and tail current decay ( $\tau_{\text{Blocker}}/\tau_{\text{Ctrl}}$ ) as a function of blocker concentration/ $K_d(0)$  (binding constant at 0 mV). **c**, BK  $\alpha$ -subunits coexpressed with construct D20A ( $\beta$ 3a N terminus appended to  $\beta$ 2 subunit, see Methods). Currents were activated with 10  $\mu$ M Ca<sup>2+</sup>, with the indicated voltage steps, and 2 (red) or 10 (blue) mM TBA. **d**, Effect of TBA concentration on tail current amplitude and time constant plotted assuming  $K_d(0)$  for TBA is 0.9 mM. **e**, BK  $\alpha$  + D20A traces for control solution and with 200 and 400  $\mu$ M bbTBA. **f**, Effect of bbTBA on tail current amplitude and time constant with a  $K_d(0)$  for bbTBA block of 6  $\mu$ M. Error bars, means  $\pm$  s.e.m. of 3–5 measurements.

blocker effects on the time course of  $\beta$ 3a-mediated currents, both for onset of inactivation (Supplementary Fig. 8b, d) and for tail currents (Supplementary Fig. 8c, d). Only the model (model 3) in which a pore blocker competitively prevents the O\*  $\rightarrow$  I\* transition predicts that blocker will reduce both tail current amplitude and duration (Fig. 4a, b).

We therefore tested the effects of the pore blocker, TBA<sup>27</sup>, on  $\beta$ 3a-mediated tail currents; both peak tail current and tail current time constant were reduced (Fig. 4c, d). These effects occur at concentrations in excess of the TBA blocking affinity for open BK channels (approximately 0.9 mM (ref. 27)), but consistent with the competitive model (Fig. 4b). Similar results were obtained with the bulky blocker, bbTBA<sup>28</sup> (Fig. 4e, f). We conclude that  $\beta$ 3a-mediated inactivation involves movement of part of the  $\beta$ 3a N terminus into a position that can also be occupied by BK pore blockers. Although it has been assumed that BK  $\beta$ -subunit N termini produce inactivation by occupying a site within the BK channel central cavity, this provides the first direct evidence of this assumption. Our results do not establish the position of the stereospecific interaction between  $\beta$ 3a peptide and the  $\alpha$ -subunit, but we suggest that the binding-site determinants are on the walls of the cavernous BK inner pore<sup>29,30</sup>.

These results demonstrate at the single molecule level that an intrinsically disordered peptide domain can bind in a stereospecific

fashion to a target protein, thereby mediating physiological effects. An important implication of the mechanism by which the  $\beta 3a$  N terminus prolongs tail current is that the  $O^*$  state—although an open, ion-conducting conformation—is unable to close until the  $\beta 3a$  peptide dissociates from the channel. We propose that not only does this intrinsically disordered peptide bind to a stereospecific site on or near the BK channel inner pore, but that this binding allosterically stabilizes the BK channel in an open conformation. The requirement for stereospecific binding in BK inactivation contrasts with *Shaker* inactivation, which is mediated largely by hydrophobic interactions of its inactivation domain with the *Shaker* pore. In each case, inactivation domain disorder may be required to facilitate movement of the N-terminal inactivation segment through narrow pathways to reach blocking positions. The differences in stereochemical requirements may be required by structural differences in the ion channel pores: the *Shaker* domain may fit more snugly in the narrower open Kv channel<sup>8</sup>, whereas in the larger BK channel inner pore<sup>30</sup> stereospecific binding may be required to achieve the affinity necessary for inhibition or to modulate channel gating.

## METHODS SUMMARY

Recordings of channel activity used inside-out patches from *Xenopus* oocytes expressing either wild-type BK channels along with wild-type or mutant  $\beta 3a$  subunits or *Shaker*-IR channels. All recordings were with symmetrical  $K^+$  solutions, with solutions of calibrated  $Ca^{2+}$  applied to the cytosolic face of BK channels. Solution control and exchange at the cytosolic face of a patch were done with a multi-barrel pipette. Currents were typically acquired at 100 kHz sampling, with filtering at 10 kHz. Single-channel records used off-line subtraction based on null-sweeps. D- and L-peptides for  $\beta 3a$  N termini (MQPFSIPVQITLQGSRRRQGR-NH<sub>2</sub>) or *Shaker* N termini MAAVAGLYGLGEDRQHRKKQ-NH<sub>2</sub> were obtained commercially. Simulations of different models of inactivation used the IChSim program. 21-mer D- and L- $\beta 3a$  peptides were dissolved at a final peptide concentration of 2 mM for acquisition of one-dimensional proton spectra and two-dimensional homonuclear phase-sensitive total correlation spectroscopy (TOCSY), NOESY and rotating-frame NOE spectroscopy (ROESY) spectra. Circular dichroism spectra were generated with a Jasco J715 after dilution of the NRM samples in water.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 20 September 2011; accepted 29 February 2012.

Published online 22 April 2012.

- Dyson, H. J. Expanding the proteome: disordered and alternatively folded proteins. *Q. Rev. Biophys.* **44**, 467–518 (2011).
- Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).
- Galea, C. A. et al. Role of intrinsic flexibility in signal transduction mediated by the cell cycle regulator, p27 Kip1. *J. Mol. Biol.* **376**, 827–838 (2008).
- Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).
- De Sancho, D. & Best, R. B. Modulation of an IDP binding mechanism and rates by helix propensity and non-native interactions: association of HIF1 $\alpha$  with CBP. *Mol. Biosyst.* **8**, 256–267 (2012).
- Hoshi, T., Zagotta, W. N. & Aldrich, R. W. Biophysical and molecular mechanisms of *Shaker* potassium channel inactivation. *Science* **250**, 533–538 (1990).
- Zagotta, W. N., Hoshi, T. & Aldrich, R. W. Restoration of inactivation in mutants of *Shaker* potassium channels by a peptide derived from ShB. *Science* **250**, 568–571 (1990).
- Zhou, M., Morais-Cabral, J. H., Mann, S. & MacKinnon, R. Potassium channel receptor site for the inactivation gate and quaternary amine inhibitors. *Nature* **411**, 657–661 (2001).
- Ruppersberg, J. P., Frank, R., Pongs, O. & Stocker, M. Cloned neuronal  $I_{K(A)}$  channels reopen during recovery from inactivation. *Nature* **353**, 657–660 (1991).
- Tseng-Crank, J., Yao, J. A., Berman, M. F. & Tseng, G. N. Functional role of the NH<sub>2</sub>-terminal cytoplasmic domain of a mammalian A-type K channel. *J. Gen. Physiol.* **102**, 1057–1083 (1993).

- Rasmusson, R. L., Wang, S., Castellino, R. C., Morales, M. J. & Strauss, H. C. The beta subunit, Kv beta 1.2, acts as a rapid open channel blocker of NH<sub>2</sub>-terminal deleted Kv1.4 alpha-subunits. *Adv. Exp. Med. Biol.* **430**, 29–37 (1997).
- Kondoh, S., Ishii, K., Nakamura, Y. & Taira, N. A mammalian transient type  $K^+$  channel, rat Kv1.4, has two potential domains that could produce rapid inactivation. *J. Biol. Chem.* **272**, 19333–19338 (1997).
- Murrell-Lagnado, R. D. & Aldrich, R. W. Interactions of amino terminal domains of *Shaker* K channels with a pore blocking site studied with synthetic peptides. *J. Gen. Physiol.* **102**, 949–975 (1993).
- Schott, M. K., Antz, C., Frank, R., Ruppersberg, J. P. & Kalbitzer, H. R. Structure of the inactivating gate from the *Shaker* voltage gated  $K^+$  channel analyzed by NMR spectroscopy. *Eur. Biophys. J.* **27**, 99–104 (1998).
- Wissmann, R. et al. NMR structure and functional characteristics of the hydrophilic N terminus of the potassium channel  $\beta$ -subunit Kv $\beta$ 1.1. *J. Biol. Chem.* **274**, 35521–35525 (1999).
- Wissmann, R. et al. Solution structure and function of the ‘tandem inactivation domain’ of the neuronal A-type potassium channel Kv1.4. *J. Biol. Chem.* **278**, 16142–16150 (2003).
- Bentrop, D., Beyermann, M., Wissmann, R. & Fakler, B. NMR structure of the ‘ball-and-chain’ domain of KCNMB2, the  $\beta_2$ -subunit of large conductance  $Ca^{2+}$ - and voltage-activated potassium channels. *J. Biol. Chem.* **276**, 42116–42121 (2001).
- Demo, S. D. & Yellen, G. The inactivation gate of the *Shaker*  $K^+$  channel behaves like an open-channel blocker. *Neuron* **7**, 743–753 (1991).
- Gonzalez, C., Lopez-Rodriguez, A., Srikumar, D., Rosenthal, J. J. & Holmgren, M. Editing of human Kv1.1 channel mRNAs disrupts binding of the N-terminus tip at the intracellular cavity. *Nature Commun.* **2**, 436, <http://dx.doi.org/10.1038/ncomms1446> (2011).
- Lingle, C. J., Zeng, X.-H., Ding, J.-P. & Xia, X.-M. Inactivation of BK channels mediated by the N-terminus of the  $\beta 3b$  auxiliary subunit involves a two-step mechanism: possible separation of binding and blockade. *J. Gen. Physiol.* **117**, 583–605 (2001).
- Zeng, X.-H., Xia, X. M. & Lingle, C. J. BK channels with  $\beta 3a$  subunits generate use-dependent slow afterhyperpolarizing currents by an inactivation-coupled mechanism. *J. Neurosci.* **27**, 4707–4715 (2007).
- Xia, X.-M., Ding, J. P. & Lingle, C. J. Molecular basis for the inactivation of  $Ca^{2+}$ - and voltage-dependent BK channels in adrenal chromaffin cells and rat insulinoma tumor cells. *J. Neurosci.* **19**, 5255–5264 (1999).
- Wallner, M., Meera, P. & Toro, L. Molecular basis of fast inactivation in voltage and  $Ca^{2+}$ -activated  $K^+$  channels: a transmembrane  $\beta$ -subunit homolog. *Proc. Natl Acad. Sci. USA* **96**, 4137–4142 (1999).
- Xia, X.-M., Ding, J.-P., Zeng, X.-H., Duan, K.-L. & Lingle, C. J. Rectification and rapid activation at low  $Ca^{2+}$  of  $Ca^{2+}$ -activated, voltage-dependent BK currents: consequences of rapid inactivation by a novel  $\beta$  subunit. *J. Neurosci.* **20**, 4890–4903 (2000).
- Prince-Carter, A. & Pfaffinger, P. J. Multiple intermediate states precede pore block during N-type inactivation of a voltage-gated potassium channel. *J. Gen. Physiol.* **134**, 15–34 (2009).
- Choi, K. L., Aldrich, R. W. & Yellen, G. Tetraethylammonium blockade distinguishes two inactivation mechanisms in voltage-activated  $K^+$  channels. *Proc. Natl Acad. Sci. USA* **88**, 5092–5095 (1991).
- Li, W. & Aldrich, R. W. Unique inner pore properties of BK channels revealed by quaternary ammonium block. *J. Gen. Physiol.* **124**, 43–57 (2004).
- Wilkins, C. M. & Aldrich, R. W. State-independent block of BK channels by an intracellular quaternary ammonium. *J. Gen. Physiol.* **128**, 347–364 (2006).
- Brelidze, T. I. & Magleby, K. L. Probing the geometry of the inner vestibule of BK channels with sugars. *J. Gen. Physiol.* **126**, 105–121 (2005).
- Zhou, Y., Xia, X. M. & Lingle, C. J. Cysteine scanning and modification reveal major differences between BK channels and Kv channels in the inner pore region. *Proc. Natl Acad. Sci. USA* **108**, 12161–12166 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by GM-081748 to C.J.L. and the Searle Scholars Program to K.H.-W. We thank E. Morrison and P. Schlesinger for assistance with dynamic light scattering measurements, and C. Frieden and K. Garai for assistance with circular dichroism spectroscopy. We thank H. Jiang, A. Scott and J. Jones for care of oocytes, and J. H. Steinbach and R. Pappu for comments on the manuscript.

**Author Contributions** V.G.-P. and X.-H.Z. designed experiments and collected and analysed data. K.H.-W. performed or supervised circular dichroism and NMR determinations. C.J.L. conceived the project, designed research, analysed data and prepared the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.J.L. ([clingle@morpheus.wustl.edu](mailto:clingle@morpheus.wustl.edu)).



## METHODS

**Constructs.** The mSlo1 construct (GenBank accession number NP\_034740) was placed in the pXMX expression construct<sup>31</sup>. The mouse  $\beta 3a$  construct (GenBank accession number NP\_001182003.1) was described in previous work<sup>32</sup>. In some experiments, construct D20A, in which the human  $\beta 3a$  N terminus was attached to a human  $\beta 2$  construct (N-terminal removed), was used<sup>21</sup>. This construct avoids the outward rectification that arises from the extracellular loop of the  $\beta 3$  subunit<sup>33</sup>. The D20A N terminus begins with 55 residues from human  $\beta 3a$  that are then appended to human  $\beta 2$  beginning at the  $\beta 2$  first transmembrane segment (h $\beta 3a$ :1–55);(h $\beta 2$ :47–235). A *Shaker-B* construct was provided by L. Jan<sup>34</sup> and residues 2–46 of the N terminus were deleted by X.-M. Xia to generate the non-inactivating *ShakerB*( $\Delta 2$ –46).

**Heterologous expression in oocytes.** Stage IV *Xenopus laevis* oocytes were used for expression of Slo1 and  $\beta 3a$  constructs. Slo1  $\alpha$  and  $\beta 3$  antisense RNA (cRNA), prepared at approximately  $1 \mu\text{g} \mu\text{l}^{-1}$ , were first diluted to 1:20 and 1:10, respectively, and the diluted solutions mixed in equal amounts. Assuming that the initial  $\beta$ -subunit RNA stock is approximately three- to fourfold higher in molar amount than the Slo1 RNA stock, the final  $\beta 3a$ : $\alpha$  molar injection ratio is approximately 6–8:1. For single-channel patches, the amount of injected RNA was reduced about 10- to 100-fold.

**Electrophysiology.** Borosilicate capillary tubes (Drummond Microcaps, 100  $\mu\text{l}$ ) were pulled to diameters resulting in access resistance of 1–2 M $\Omega$ , coated with Sylgard (Sylgard 184, Dow Chemical Corp.) and fire-polished. Currents were recorded in the inside-out configuration<sup>35</sup> using an Axopatch 200 amplifier (Molecular Devices) and the Clampex program from the pClamp software package (Molecular Devices).

Patches with gigaohm seals were formed in normal frog Ringer (in mM: 115 NaCl, 2.5 KCl, 1.8 CaCl<sub>2</sub>, 10 HEPES, pH 7.4) and then, after excision, moved into flowing test solutions to control the solution bathing the membrane face. The pipette/extracellular solution was (in mM): 140 K-methanesulphonate, 20 KOH, 10 HEPES, 2 MgCl<sub>2</sub>, pH 7.0. Test solutions bathing the cytoplasmic face of the patch membrane contained (in mM) 140 K-methanesulphonate, 20 KOH, 10 HEPES, with pH adjusted to 7.0 KOH. HEDTA (5 mM) was used for 10  $\mu\text{M}$  Ca<sup>2+</sup> solutions and 5 mM EGTA for 0  $\mu\text{M}$  Ca<sup>2+</sup>. The 10  $\mu\text{M}$  Ca<sup>2+</sup> solution was titrated to appropriate pCa with Ca-MES and calibrated against solutions of defined Ca<sup>2+</sup> concentrations (World Precision Instruments) using a Ca<sup>2+</sup>-sensitive electrode. Solutions bathing the cytosolic face of the membrane were controlled by a local application system containing up to six independent lines. Experiments were at room temperature (22–24 °C) and chemicals for solution preparation were from Sigma.

**Data analysis.** Analysis of current recordings and simulated currents was accomplished either with Clampfit (Molecular Devices) or with programs written in this

laboratory. Single-channel traces were first processed using digital subtraction of leak and capacity currents defined from traces lacking any channel openings. For ensemble averaging of *Shaker* single-channel openings, each sweep was first filtered at 1 kHz (acquisition at 50 kHz with filtering at 5 kHz).

**Simulation of currents.** Current simulations were accomplished with the IChSim program (<http://www.ifisica.uaslp.mx/~jadsc/ichsim.htm>) developed at the Physics Institute of the University of San Luis Potosi, Mexico, by J. A. De Santiago Castillo.

**Solution NMR.** Peptides (21-mer D- and L- $\beta 3a$ ) were dissolved in 20 mM potassium phosphate, 20 mM sodium chloride, pH 7.0, 10% D<sub>2</sub>O with a final peptide concentration of 2 mM. One-dimensional proton spectra and two-dimensional homonuclear phase-sensitive TOCSY, NOESY and ROESY spectra were recorded on a 600 MHz Bruker Avance III spectrometer equipped with QCI cryoprobe. NOESY and ROESY spectra were acquired with a mixing time of 300 ms, TOCSY with an 80 ms mixing time. All two-dimensional homonuclear experiments were acquired with 512 points in the indirect dimension, and 16 scans per increment. Spectra were processed with NMRpipe and displayed using NMRViewJ (two-dimensional spectra) or IgorPro (one-dimensional spectra).

**Circular dichroism spectra.** The NMR samples were diluted 40-fold in water for measurement of circular dichroism spectra using a Jasco J715.

**Peptides.** L- and D-amino-acid  $\beta 3a$  peptides and L- and D-20-mer *ShakerB* ball peptides were custom synthesized by Biomolecules Midwest with an amidated carboxy (C) terminus and NH<sub>2</sub>-N terminus, and purified by high-performance liquid chromatography to over 95% purity. Both 20 and 21 L-amino-acid versions of the  $\beta 3a$  peptides were used in separate experiments. The 21-mer  $\beta 3a$  peptide contains an additional R in position 21 and shows a faster forward rate of block than the 20-mer. The sequence of the 20-mer *ShakerB* ball peptides was MAAVAGLYGLGEDRQHRKKQ-NH<sub>2</sub>, identical to +2 *ShakerB* peptides used in previous work<sup>6,13</sup>.

31. Tang, Q., Zeng, X.-H. & Lingle, C. J. Closed channel block of BK potassium channels by bbTBA requires partial activation. *J. Gen. Physiol.* **134**, 409–436 (2009).
32. Zeng, X., Xia, X. M. & Lingle, C. J. Species-specific differences among KCNMB3 BK  $\beta 3$  auxiliary subunits: some  $\beta 3$  variants may be primate-specific subunits. *J. Gen. Physiol.* **132**, 115–129 (2008).
33. Zeng, X.-H., Xia, X.-M. & Lingle, C. J. Redox-sensitive extracellular gates formed by auxiliary  $\beta$  subunits of calcium-activated potassium channels. *Nature Struct. Biol.* **10**, 448–454 (2003).
34. Timpe, L. C. *et al.* Expression of functional potassium channels from *Shaker* cDNA in *Xenopus* oocytes. *Nature* **331**, 143–145 (1988).
35. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Arch.* **391**, 85–100 (1981).

# CAREERS

**ACADEMIA** Men receive informal benefits more frequently than women **p.139**

**PROFESSIONAL DEVELOPMENT** More support needed for diverse career options **p.139**

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)



RESEARCH

## Uncovering misconduct

*Cases of scientific wrongdoing seem to be rising. But when should researchers blow the whistle?*

BY VIRGINIA GEWIN

Biostatisticians Keith Baggerly and Kevin Coombes, like many, were intrigued by claims of personalized chemotherapy treatments by geneticist Anil Potti in 2006. Then at Duke University in Durham, North Carolina, Potti published results indicating that gene expression signatures could identify which chemotherapy drug could best treat lung or breast cancer — results that led to the setup of three clinical trials. But Baggerly and Coombes quickly found something amiss in the data. What began as concerns over apparent errors, including mislabelled samples and mismatched gene names, eventually snowballed into one of the most notorious cases of scientific misconduct in the United States in recent years.

During some 1,500 hours of work over four years, Baggerly and Coombes, both of the University of Texas MD Anderson Cancer Center in Houston, repeatedly showed that Potti's findings did not match the raw data. They analysed the data, had conversations with Potti and his

supervisor, alerted the US National Cancer Institute to the likely mistakes and contacted the editors of the journal publishing Potti's work.

Repeated enquiries and complaints by Baggerly and Coombes led senior officials at the University of Texas to advise them to drop what was starting to look like a vendetta. "We were focused on the fact that the data used to justify clinical trials were wrong; we thought that should be enough," says Baggerly. "How the data got in this shape was not our immediate concern."

Their objections were finally proved valid. Six years on, ten papers by Potti have been retracted, and the clinical trials were halted eventually. Baggerly and Coombes say that their persistence was down to their obligation to scientific ethics and the consequences of a clinical trial based on incorrect data.

Gauging the amount of misconduct in science is very difficult, but last year there were 381 journal retraction notices — up from 22 in 2001 — according to the Thomson Reuters database Web of Knowledge. Indeed, 2011 was dubbed the "year of the retraction"

by the blog Retraction Watch. Last year also saw 13 misconduct rulings by the US Office of Research Integrity (ORI) in Rockville, Maryland, which oversees misconduct investigations and publishes the findings on its website.

Some reports suggest cases of misconduct may be more prevalent than previously suspected. In January, a survey<sup>1</sup> in the *British Medical Journal* found that of the 2,782 doctors and academics that responded, 13% had first-hand knowledge of misconduct. These findings mirror a 2009 meta-analysis of misconduct surveys conducted by Daniele Fanelli from the University of Edinburgh, UK, which found that almost 2% of scientists admitted to having fabricated, falsified or modified data or results at least once, and 14% knew of fabrication or falsification by colleagues<sup>2</sup>. "Yet consistently, almost no one reports misconduct to the proper authorities," says Fanelli. According to his analysis, five studies asked survey respondents if they had taken any action to correct or prevent misconduct. Only about half of the alleged cases that were reported resulted in any action — and even then, this amounted

B. MELLOR

► mainly to informal confrontations or discussions with colleagues.

According to a 2010 survey<sup>3</sup> on the response of researchers to wrongdoing, 63% of the 2,193 respondents said that they had intervened but that most action was informal — discussing concerns with a supervisor or questioning the suspect behaviour, rather than lodging a formal complaint.

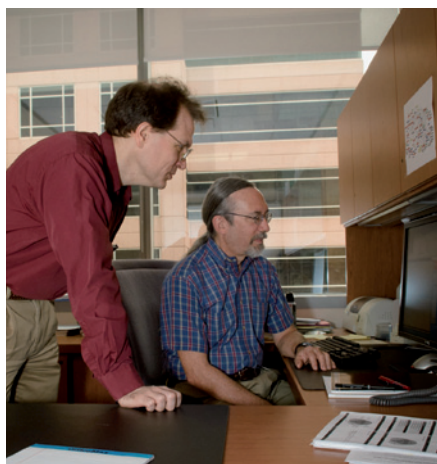
Would-be whistleblowers face tough decisions about whether their concerns are the result of misconduct, scientific disagreement or simply an honest error. Knowing where or how to lodge complaints can also be difficult, given that misconduct investigations are pursued confidentially. From identifying the proper place to report allegations to understanding how an investigation will unfold, there are a number of considerations, policies and procedures that would-be whistleblowers need to know.

### TO ALLEGE OR NOT?

The first step towards determining whether to air integrity concerns is to understand what is and what is not scientific misconduct. Misconduct is not simply bad behaviour; it is the falsification, fabrication or plagiarism of results. Honest errors, differences in the interpretation of results, authorship disputes, sexual harassment or threatening language are issues of concern, but are not misconduct. At the core of misconduct is intent. “There is not a finding of misconduct unless it can be proven that the person acted with intention or was seriously reckless,” says Mark Barnes, chief research compliance officer at Harvard University in Cambridge, Massachusetts.

Plagiarism cases have become fairly easy to identify — in part because of software tools such as Turnitin and iThenticate. Increasingly common are allegations of image manipulations (for example, falsely labelling the protein bands on a gel), which can also be identified quickly with software.

Even with a clear definition of misconduct, grey areas can still arise. Cherry-picking statistical methods to obtain desired results or omitting outlier points on a graph can fall under the rubric of the honest difference in the interpretation of results. Evolutionary biologist Marc Hauser resigned from Harvard last year after he was found solely responsible for eight instances of scientific misconduct, most notably in a paper showing that cotton-top tamarin monkeys could learn simple rule-like patterns. The paper was retracted in 2010. A former trainee of Hauser’s, who prefers to remain anonymous, did not agree with the final findings of the investigation. “I can see data getting overlooked or favourably interpreted, but I would be surprised if data had been intentionally changed,” says the source. Several of the experiments in question have since been replicated by Hauser and colleagues and the results published, highlighting the fine line between a



Keith Baggerly (left) and Kevin Coombes felt they had an obligation to uphold scientific ethics.

lack of data integrity and misconduct.

Researchers who are concerned about the conduct of a colleague should make sure that they understand the nature of the research, says David Resnik, a bioethicist and chair of the Institutional Review Board for the National Institute of Environmental Health Sciences in Research Triangle Park, North Carolina. If someone has concerns about whether research methods are appropriate, Resnik says that they should consult a colleague who is familiar with the accepted scientific norms of that discipline.

If the concerns persist, the next step is to decide how to lodge the complaint — either anonymously using telephone hotlines found in many institutes, or in person. The decision can be difficult. Hotline calls may protect the

whistleblower, but may not lead to a misconduct investigation. “Anonymous reports may not allow the university to gather enough credible information to proceed with an investigation,” says Jan Allen, associate dean for PhD programmes at Columbia University in New York.

In the United States, legislation exists to protect whistleblowers from retaliation, and universities do their best to maintain confidentiality throughout an investigation. “Protections are in place to keep complainants anonymous, but confidentiality can’t be absolutely guaranteed — for example, if university policy is being violated, we have to report it,” says Allen. In some cases, the university can move complainants to a different lab. “But it is hard to deal with lesser acts of retaliation, such as the refusal to write a letter of recommendation,” admits David Wright, a former research integrity officer (RIO) at Michigan State University in East Lansing and the director of the ORI. Those contemplating blowing the whistle should consider the university’s policies on misconduct.

### SEEKING COUNSEL

In the United States, a university’s RIO is hired by that institution to handle allegations of research misconduct. But faculty and staff are often not aware that they exist and thus do not seek their help. Furthermore, not all researchers will receive the same assistance, and not all countries have these measures (see ‘International standards’). A study that interviewed 79 RIOs<sup>4</sup> found that they were “not uniformly well prepared” to handle allegations of scientific

## INTERNATIONAL STANDARDS

### Oversight measures in other countries

In the United States, whistleblowers are offered protection and are, increasingly, encouraged to report cases of misconduct. But that is not the case everywhere.

Without such protections, academics in the United Kingdom who want to report misconduct may find themselves suspended or the focus of retribution. As such, the only sure-fire way to get allegations heard is by airing them to the press. And UK libel laws make filing a suit against whistleblowers relatively easy. (However, proposals to reform libel laws are included in a draft defamation bill before Parliament.) “I tell people to think hard about whether you want to blow the whistle,” says Peter Wilmshurst, a cardiologist and well-known whistleblower in the United Kingdom who reported 25 people he believed were conducting fraudulent research to the UK General Medical Council. “You have to decide whether you are willing to run the risk of ruining your career and

never working again.”

Fiona Godlee, editor-in-chief of the *British Medical Journal*, wants the United Kingdom to establish similar oversight measures to the United States, possibly by giving the UK Research Integrity Office, currently an advisory body, some regulatory power. Godlee would also like to see a policy that requires funding agencies to give money only to institutions that have a code of conduct, a commitment to investigate allegations and an appointed person for research integrity. “We want to put a preventative culture in place,” she says.

Few international standards exist. Laura Marin, science officer for the European Science Foundation based in Strasbourg, France, is working on the implementation of a Europe-wide code of conduct for research integrity. “Misconduct is a sensitive issue — which is why, at this moment, there is not a clear vision to get to legislation.” **VG.**



misconduct, and that 54% of RIOs had never called ORI to report misconduct. Wright says that cultures vary greatly by campus, which is one reason why the ORI boot camp, launched in 2007, offers ongoing formal training to RIOs. Still, there is some indication that if whistleblowers are dissatisfied with their institution's response, they may have to contact the ORI themselves. "When the ORI receives allegations that are substantial, we request that the institution move immediately to an inquiry," says John Dahlberg, director of the division of investigative oversight at the ORI. UK universities may also have designated RIOs, if funders adhere to calls made in February, following a research misconduct meeting organized by the *British Medical Journal* and the Committee on Publication Ethics.

Specificity is a key component of any evidence used to substantiate an allegation, says Dahlberg. Resnik says "anytime you go forward, you need documentation to back up what you say so the allegations are not tossed out." Whistleblowers should never file a formal complaint on the basis of a rumour or information gained from a third party, says Gerald Koocher, associate provost at Simmons College in Boston, Massachusetts. "If you don't have a smoking gun, at least have a gun," he says. Ideally, whistleblowers will be able to describe the nature and whereabouts of any additional evidence that may support the allegation, says Wright.

Once allegations are made, there is a danger that data sets could become adulterated or vanish, says Barnes. The ORI requires institutions that have received credible allegations to seize the computer, e-mails or data that may be used as evidence, to prevent this happening.

However, tipping off the perpetrator is a valid concern in a lab where people work closely together. Colleagues are likely to discuss their suspicions before making a formal allegation, creating an opportunity for the perpetrator to tamper with evidence. If researchers really believe there may be misconduct, they should either make a copy of the raw data before suspicions are aired or go straight to the authorities with their suspicions, says Barnes. The worst thing complainants can do is convince themselves that they are the prosecutor who needs to build a case against the suspect, says Wright. Complainants are wise to simply give any evidence to an impartial investigator; otherwise, their motives could be called into question.

## INVESTIGATION UNFOLDS

If an allegation is deemed to have merit, a university committee starts an inquiry to review the evidence supporting the allegation and to decide whether a formal investigation is necessary. Inquiries found to have sufficient evidence will often then lead to the formation of a new committee to undertake the investigation.

In Wright's experience, at a research-intensive institution it is not uncommon to have ten significant allegations made in a year. Of those, only about six will go to an inquiry; two may evolve into investigations, and only one or none at all will result in findings of misconduct.

Complainants should also realize that investigations can go on for a year or more. During an investigation, Barnes suggests carrying on as normal. For example, during the Hauser investigation, a former colleague says that although senior members of the lab were aware of the investigation, it was hardly ever discussed. But Koocher adds that complainants are wise to document everything they witness during the course of the investigation.

## PREVENTING FRAUD

High-profile misconduct cases such as those of Hauser and Potti mean that data undergo increasing scrutiny by university administrators. In the wake of the Potti case, Duke University is planning a 'data lockbox', essentially an electronic means to track who has handled data and files, and the changes they have made. The university also plans to embed biostatisticians within clinical research groups to help prevent against inadvertent errors in data analysis.

"Part of the problem with complex data sets inherent to today's science is that you can't pick them up and know instantly that something is fishy," says Sally Kornbluth, vice dean for basic sciences at Duke's School of Medicine. Resnik notes, for example, that genome-wide association studies have been controversial because of the evolving statistical methods that people are using. Baggerly advocates more open sharing of data-analysis methods. "The main things that we were stymied by was simply trying to get the raw data and the code used to perform Potti's analyses," he says.

Being vigilant in cases of apparent fraud or misconduct not only corrects the record, but saves others from wasting time, effort and money. For graduate students or postdocs, deciding whether to publicly question the practices of their colleagues can be tough. But Koocher reminds junior scientists that they are often the first to take the fall if something fails or proves unreliable in the lab. "In cases where misconduct is suspected, it's way better, and smarter, to take action that is self-protective," he says, "rather than risk getting any of the blame." ■

*Virginia Gewin is a freelance writer based in Portland, Oregon.*

1. Tavare, A. *Br. Med. J.* **344**, e377 (2012).
2. Fanelli, D. *PLoS ONE* **4**, e5738 (2009).
3. Koocher, G. P. & Keith-Spiegel, P. *Nature* **466**, 438–440 (2010).
4. Bonito, A. J., Titus, S. L. & Wright, D. E. *Sci. Eng. Ethics* <http://dx.doi.org/10.1007/s11948-011-9274-2> (2011).

## ACADEMIA

### Women miss out

Female academics across all fields are less likely than their male colleagues to receive bonuses, according to a study of employees at a large, unnamed Canadian university (C. Doucet *et al. Ind. Relat.* **67**, 51–75; 2012). The discrepancy may be because female faculty members have fewer networking connections and less knowledge about bonuses than men, suggests Christine Doucet, a sociologist at the University of Montreal, Canada, and co-author of the article, which used data on some 1,900 faculty members. Those who lack institutional networks should seek out information about informal benefits, she advises. If universities followed more formal compensation practices, rather than relying on informal discretion, equity would improve, she notes.

## PROFESSIONAL DEVELOPMENT

### Career-path support

US universities, federal policy-makers and employers must coordinate their efforts to improve the career paths of postgraduates, according to a report by the US Council of Graduate Schools and the Educational Testing Service, both in Washington DC. *Pathways Through Graduate School and Into Careers* calls for universities to offer professional-skills development training, information on non-academic careers and tracking of career outcomes. More employers need to offer student-training programmes such as internships, help to foster graduate programmes tailored to workforce needs and support employees' graduate study. The report also calls for US visa policies that help to retain international talent.

## PHARMACEUTICAL INDUSTRY

### UK placements down

Training for academics by UK drugmakers declined from 2007 to 2011, finds a survey by the London-based Association of the British Pharmaceutical Industry. The number of research-training placements fell owing, in part, to outsourcing and site closures. The number of industry postdoc positions dropped by more than 12%, and posts for undergraduates decreased by half. But support for PhD students is up because companies are moving towards funding for four years, rather than three, to offer broader training. Association spokeswoman Louise Leong notes that industry training schemes help to tailor the workforce, which facilitates job placement.

misconduct, and that 54% of RIOs had never called ORI to report misconduct. Wright says that cultures vary greatly by campus, which is one reason why the ORI boot camp, launched in 2007, offers ongoing formal training to RIOs. Still, there is some indication that if whistleblowers are dissatisfied with their institution's response, they may have to contact the ORI themselves. "When the ORI receives allegations that are substantial, we request that the institution move immediately to an inquiry," says John Dahlberg, director of the division of investigative oversight at the ORI. UK universities may also have designated RIOs, if funders adhere to calls made in February, following a research misconduct meeting organized by the *British Medical Journal* and the Committee on Publication Ethics.

Specificity is a key component of any evidence used to substantiate an allegation, says Dahlberg. Resnik says "anytime you go forward, you need documentation to back up what you say so the allegations are not tossed out." Whistleblowers should never file a formal complaint on the basis of a rumour or information gained from a third party, says Gerald Koocher, associate provost at Simmons College in Boston, Massachusetts. "If you don't have a smoking gun, at least have a gun," he says. Ideally, whistleblowers will be able to describe the nature and whereabouts of any additional evidence that may support the allegation, says Wright.

Once allegations are made, there is a danger that data sets could become adulterated or vanish, says Barnes. The ORI requires institutions that have received credible allegations to seize the computer, e-mails or data that may be used as evidence, to prevent this happening.

However, tipping off the perpetrator is a valid concern in a lab where people work closely together. Colleagues are likely to discuss their suspicions before making a formal allegation, creating an opportunity for the perpetrator to tamper with evidence. If researchers really believe there may be misconduct, they should either make a copy of the raw data before suspicions are aired or go straight to the authorities with their suspicions, says Barnes. The worst thing complainants can do is convince themselves that they are the prosecutor who needs to build a case against the suspect, says Wright. Complainants are wise to simply give any evidence to an impartial investigator; otherwise, their motives could be called into question.

## INVESTIGATION UNFOLDS

If an allegation is deemed to have merit, a university committee starts an inquiry to review the evidence supporting the allegation and to decide whether a formal investigation is necessary. Inquiries found to have sufficient evidence will often then lead to the formation of a new committee to undertake the investigation.

In Wright's experience, at a research-intensive institution it is not uncommon to have ten significant allegations made in a year. Of those, only about six will go to an inquiry; two may evolve into investigations, and only one or none at all will result in findings of misconduct.

Complainants should also realize that investigations can go on for a year or more. During an investigation, Barnes suggests carrying on as normal. For example, during the Hauser investigation, a former colleague says that although senior members of the lab were aware of the investigation, it was hardly ever discussed. But Koocher adds that complainants are wise to document everything they witness during the course of the investigation.

## PREVENTING FRAUD

High-profile misconduct cases such as those of Hauser and Potti mean that data undergo increasing scrutiny by university administrators. In the wake of the Potti case, Duke University is planning a 'data lockbox', essentially an electronic means to track who has handled data and files, and the changes they have made. The university also plans to embed biostatisticians within clinical research groups to help prevent against inadvertent errors in data analysis.

"Part of the problem with complex data sets inherent to today's science is that you can't pick them up and know instantly that something is fishy," says Sally Kornbluth, vice dean for basic sciences at Duke's School of Medicine. Resnik notes, for example, that genome-wide association studies have been controversial because of the evolving statistical methods that people are using. Baggerly advocates more open sharing of data-analysis methods. "The main things that we were stymied by was simply trying to get the raw data and the code used to perform Potti's analyses," he says.

Being vigilant in cases of apparent fraud or misconduct not only corrects the record, but saves others from wasting time, effort and money. For graduate students or postdocs, deciding whether to publicly question the practices of their colleagues can be tough. But Koocher reminds junior scientists that they are often the first to take the fall if something fails or proves unreliable in the lab. "In cases where misconduct is suspected, it's way better, and smarter, to take action that is self-protective," he says, "rather than risk getting any of the blame." ■

*Virginia Gewin is a freelance writer based in Portland, Oregon.*

1. Tavaré, A. *Br. Med. J.* **344**, e377 (2012).
2. Fanelli, D. *PLoS ONE* **4**, e5738 (2009).
3. Koocher, G. P. & Keith-Spiegel, P. *Nature* **466**, 438–440 (2010).
4. Bonito, A. J., Titus, S. L. & Wright, D. E. *Sci. Eng. Ethics* <http://dx.doi.org/10.1007/s11948-011-9274-2> (2011).

## ACADEMIA

### Women miss out

Female academics across all fields are less likely than their male colleagues to receive bonuses, according to a study of employees at a large, unnamed Canadian university (C. Doucet *et al. Ind. Relat.* **67**, 51–75; 2012). The discrepancy may be because female faculty members have fewer networking connections and less knowledge about bonuses than men, suggests Christine Doucet, a sociologist at the University of Montreal, Canada, and co-author of the article, which used data on some 1,900 faculty members. Those who lack institutional networks should seek out information about informal benefits, she advises. If universities followed more formal compensation practices, rather than relying on informal discretion, equity would improve, she notes.

## PROFESSIONAL DEVELOPMENT

### Career-path support

US universities, federal policy-makers and employers must coordinate their efforts to improve the career paths of postgraduates, according to a report by the US Council of Graduate Schools and the Educational Testing Service, both in Washington DC. *Pathways Through Graduate School and Into Careers* calls for universities to offer professional-skills development training, information on non-academic careers and tracking of career outcomes. More employers need to offer student-training programmes such as internships, help to foster graduate programmes tailored to workforce needs and support employees' graduate study. The report also calls for US visa policies that help to retain international talent.

## PHARMACEUTICAL INDUSTRY

### UK placements down

Training for academics by UK drugmakers declined from 2007 to 2011, finds a survey by the London-based Association of the British Pharmaceutical Industry. The number of research-training placements fell owing, in part, to outsourcing and site closures. The number of industry postdoc positions dropped by more than 12%, and posts for undergraduates decreased by half. But support for PhD students is up because companies are moving towards funding for four years, rather than three, to offer broader training. Association spokeswoman Louise Leong notes that industry training schemes help to tailor the workforce, which facilitates job placement.

# LOYALTY BEYOND SEASONS

*Temporal reproductive isolation in Homo sapiens.*

BY MOHSEN H. DARABI

I started to have feelings for her in the October. We were in the same class. I presented her with a red rose in an attempt to express my love for her, but she refused to accept it. A month later, I asked my aunt to talk to her after class on my behalf. That didn't go well either. "I'm busy with my courses," she told my aunt. "I won't have much time for other matters."

As the climate got colder, my feelings for her began to fade away. As December turned into the January of a new year, I often found myself thinking that the very idea of a relationship with her was ridiculous. But spring came, and with it a rise in temperature. She started to stare at me in class. Sometimes she smiled at me and blushed. It was in April that she finally came up to me. "Let's have lunch together," she said. "I will be in the dining hall at 12:30."

I agreed to meet her, but in my heart I was not keen — my instinct told me that she wasn't for me, so I skipped the date and ate lunch alone. She remained persistent in her desire, but I had lost my affection for her. She sent me a message reminding me of my attentions in the previous autumn. She mentioned the red rose and my aunt. But try as I might, I couldn't get excited about being with her any more. The season had changed.

Spring turned into summer. Her feelings towards me lessened slowly from about June onwards, but I thought that she still liked me. Autumn came, taking summer's heat. To my surprise, I began loving her again — exactly as I had the year before. This time, I decided to see her father and ask for his daughter's hand in marriage. Finding his address was a problem. I was ashamed to ask her for it directly, given that I had spurned her attentions back in the spring. I found it, eventually, and I met him that November.

"It depends on my daughter's opinion," her father said. "If she agrees, I will have no objection."

"Might I call you tomorrow to know her answer?" I asked.

"No, I'll call you myself and let you know,"

he insisted. I had no choice but to agree.

He didn't call. Eleven days elapsed — there was still no sign of his number on my phone. I finally decided to talk to her directly. I don't want to go into everything she told me then, but, in summary, she had lost interest in me. I remembered the rose, and how she had refused it. I remembered my aunt's unsuccessful intervention.

you will excuse me..."

"No Mohsen, wait —" she interrupted. "I made a mistake in refusing you. Forget about the past. Let's be friends now."

"I'm afraid I can't, not any more," I said. "I'm sorry." I walked away but her next words brought me to a dead stop.

"But you spoke to my father," she pleaded, "and even before that you wanted me, isn't that true?" I turned and saw tears in her eyes for the first time. Slowly but eagerly, she stepped closer and very faintly, she said: "Kiss me, Mohsen. I'm home." I dove into the pacific depths of her desperate heart and, although I wasn't attracted to her at that time of the year, I decided to kiss her.

I decided to, not because I was attracted to her, but because I am a human. I can think and remember things. I remembered how I had felt about her each autumn. She was certainly feeling the same about me each spring. I kissed her quivering lip because I cared about her and didn't want to see her upset. On my part, this had nothing to do with physiological love. Commitment following conscious thought is a characteristic of human behaviour.

As we kissed, I cursed Mother Nature for programming our mating seasons differently. I was cursing because seasonal changes altered our hormonal

balances in different ways.

The fact is that she gets passionate about me when temperatures rise in spring, but I am attracted to her when temperatures fall in autumn. This means that we will never feel passionate about each other simultaneously. But as two humans, we have the potential to think about these facts and respect each other's physiologies. This is when true love shows itself; when two people love each other by understanding and respecting each other.

By the autumn of 2011, she too had become aware of these concepts and promised to care about me likewise.

It is now 7 December 2011. So far, our commitment lives on. ■

**Mohsen H. Darabi** is a 21-year-old male human. He dreams of establishing a reindeer farm in Persia and a camel farm in Sweden.



With a heavy heart, I walked away.

The year turned. Winter came once again, and spring followed. In April, she started looking at me lustfully again, just as she had the year before. I became frustrated, and then angry, as it was now obvious that I had no sense of romance in me at that time of the year and she was exactly the opposite.

She confronted me when I was alone, somewhere quiet, but was too embarrassed to say anything more than: "Hello."

I broke the awkward silence with polite formalities: I asked how her father was, or some such. She looked into my eyes and smiled. "He's fine and he sometimes asks me how you're doing."

"Please tell him I'm well and that I send him my best regards," I replied, stuffily. "Now if

➔ **NATURE.COM**  
Follow Futures on  
Facebook at:  
[go.nature.com/mtoodm](http://go.nature.com/mtoodm)



# Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets

Masaki Imai<sup>1</sup>, Tokiko Watanabe<sup>1,2</sup>, Masato Hatta<sup>1</sup>, Subash C. Das<sup>1</sup>, Makoto Ozawa<sup>1,3</sup>, Kyoko Shinya<sup>4</sup>, Gongxun Zhong<sup>1</sup>, Anthony Hanson<sup>1</sup>, Hiroaki Katsura<sup>5</sup>, Shinji Watanabe<sup>1,2</sup>, Chengjun Li<sup>1</sup>, Eiryo Kawakami<sup>2</sup>, Shinya Yamada<sup>5</sup>, Maki Kiso<sup>5</sup>, Yasuo Suzuki<sup>6</sup>, Eileen A. Maher<sup>1</sup>, Gabriele Neumann<sup>1</sup> & Yoshihiro Kawaoka<sup>1,2,3,5</sup>

Highly pathogenic avian H5N1 influenza A viruses occasionally infect humans, but currently do not transmit efficiently among humans. The viral haemagglutinin (HA) protein is a known host-range determinant as it mediates virus binding to host-specific cellular receptors<sup>1–3</sup>. Here we assess the molecular changes in HA that would allow a virus possessing subtype H5 HA to be transmissible among mammals. We identified a reassortant H5 HA/H1N1 virus—comprising H5 HA (from an H5N1 virus) with four mutations and the remaining seven gene segments from a 2009 pandemic H1N1 virus—that was capable of droplet transmission in a ferret model. The transmissible H5 reassortant virus preferentially recognized human-type receptors, replicated efficiently in ferrets, caused lung lesions and weight loss, but was not highly pathogenic and did not cause mortality. These results indicate that H5 HA can convert to an HA that supports efficient viral transmission in mammals; however, we do not know whether the four mutations in the H5 HA identified here would render a wholly avian H5N1 virus transmissible. The genetic origin of the remaining seven viral gene segments may also critically contribute to transmissibility in mammals. Nevertheless, as H5N1 viruses continue to evolve and infect humans, receptor-binding variants of H5N1 viruses with pandemic potential, including avian–human reassortant viruses as tested here, may emerge. Our findings emphasize the need to prepare for potential pandemics caused by influenza viruses possessing H5 HA, and will help individuals conducting surveillance in regions with circulating H5N1 viruses to recognize key residues that predict the pandemic potential of isolates, which will inform the development, production and distribution of effective countermeasures.

Although H5N1 viruses continue to cause outbreaks in poultry and there are cases of human infection in Indonesia, Vietnam, Egypt and elsewhere ([http://www.who.int/influenza/human\\_animal\\_interface/H5N1\\_cumulative\\_table\\_archives/en/index.html](http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/index.html)), they have not acquired the ability to cause human-to-human transmission. Investment in H5N1 vaccines has therefore been questioned. However, because humans lack immunity to influenza viruses possessing an H5 HA, the emergence of a transmissible H5-HA-possessing virus would probably cause a pandemic. To prepare better for such a scenario, it is critical that we understand the molecular changes that may render H5-HA-possessing viruses transmissible in mammals. Such knowledge would allow us to monitor circulating or newly emerging variants for their pandemic potential, focus eradication efforts on viruses that already have acquired subsets of molecular changes critical for transmission in mammals, stockpile antiviral compounds in regions where such viruses circulate, and initiate vaccine generation and large-scale production

before a pandemic. Therefore, we studied the molecular features that would render H5-HA-possessing viruses transmissible in mammals.

Previous studies suggested that HA has a major role in host-range restriction of influenza A viruses<sup>1–3</sup>. The HA of human isolates preferentially recognizes sialic acid linked to galactose by  $\alpha$ 2,6-linkages (Sia $\alpha$ 2,6Gal), whereas the HA of avian isolates preferentially recognizes sialic acid linked to galactose by  $\alpha$ 2,3-linkages (Sia $\alpha$ 2,3Gal)<sup>3</sup>. A small number of avian H5N1 viruses isolated from humans show limited binding to human-type receptors, a property conferred by several amino acid changes in HA<sup>4–9</sup>. None of the H5N1 viruses tested transmitted efficiently in a ferret model<sup>10–13</sup>, although, while our paper was under review, one study<sup>14</sup> reported that a virus with a mutant H5 HA and a neuraminidase (NA) of a human virus in the H5N1 virus background caused respiratory droplet transmission in one of two contact ferrets.

To identify novel mutations in avian H5 HAs that confer human-type receptor-binding preference, we introduced random mutations into the globular head (amino acids 120–259 (H3 numbering), which includes the receptor-binding pocket) of A/Vietnam/1203/2004 (H5N1; VN1203) HA (Supplementary Fig. 1). Although this virus was isolated from a human, its HA retains avian-type receptor-binding properties<sup>6,15</sup>. We also replaced the multibasic HA cleavage sequence with a non-virulent-type cleavage sequence, allowing us to perform studies in biosafety level 2 containment ([http://www.who.int/csr/resources/publications/influenza/influenzaRMD2003\\_5.pdf](http://www.who.int/csr/resources/publications/influenza/influenzaRMD2003_5.pdf)). The mutated polymerase chain reaction (PCR) products were cloned into RNA polymerase I plasmids<sup>16</sup> containing the VN1203 HA complementary DNA, which resulted in *Escherichia coli* libraries representing the randomly generated HA variants. Sequence analysis of 48 randomly selected clones indicated an average of 1.0 amino acid changes per HA globular head (data not shown). To generate an H5N1 virus library, plasmids for the synthesis of the mutated HA gene and the unmodified NA gene of VN1203 were transfected into human embryonic kidney (293T) cells together with plasmids for the synthesis of the six remaining viral genes of A/Puerto Rico/8/34 (H1N1; PR8), a laboratory-adapted human influenza A virus.

Turkey red blood cells (TRBCs; which possess both Sia $\alpha$ 2,6Gal and Sia $\alpha$ 2,3Gal on their surface (data not shown)) were treated with *Salmonella enterica* serovar Typhimurium LT2 sialidase, which preferentially removes  $\alpha$ 2,3-linked sialic acid (that is, avian-type receptors), creating TRBCs that predominantly possess Sia $\alpha$ 2,6Gal on the cell surface (Sia $\alpha$ 2,6-TRBCs; Supplementary Fig. 2). The virus library was then adsorbed to Sia $\alpha$ 2,6-TRBCs at 4 °C and extensively washed to remove nonspecifically or weakly bound viruses. Bound viruses were eluted by incubation at 37 °C for 30 min, and then diluted to approximately ~0.5 viruses per well (on the basis of a pilot experiment that

<sup>1</sup>Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53711, USA. <sup>2</sup>ERATO Infection-Induced Host Responses Project, Saitama 332-0012, Japan. <sup>3</sup>Department of Special Pathogens, International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. <sup>4</sup>Department of Microbiology and Infectious Diseases, Kobe University, Hyogo 650-0017, Japan. <sup>5</sup>Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. <sup>6</sup>Health Science Hills, College of Life and Health Sciences, Chubu University, Kasugai, Aichi 487-8501, Japan.

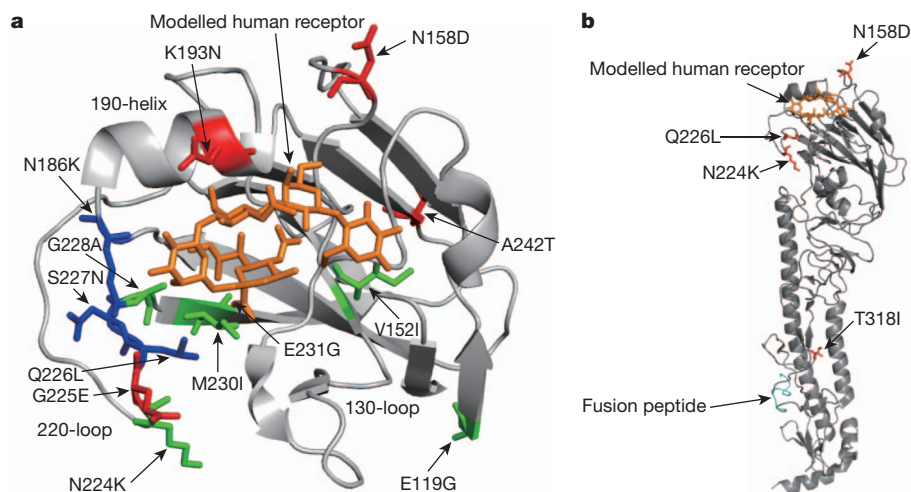
assessed the approximate number of eluted viruses). We screened one-third of the library (that is,  $2.1 \times 10^6$  viruses) in three separate selection experiments (that is,  $0.7 \times 10^6$  viruses per experiment) and isolated 370 viruses that bound to Sia $\alpha$ 2,6-TRBCs (Supplementary Fig. 2). Individual viruses were then grown in Madin-Darby canine kidney (MDCK) cells modified to overexpress Sia $\alpha$ 2,6Gal (AX4 cells<sup>17</sup>), and screened again for their ability to agglutinate Sia $\alpha$ 2,6-TRBCs (Supplementary Fig. 2). The parental control virus (designated VN1203/PR8) with avian-type receptor-binding specificity agglutinated untreated TRBCs (which possess both human- and avian-type receptors on their surface), but not TRBCs possessing predominantly human-type receptors (Sia $\alpha$ 2,6-TRBCs; Supplementary Table 1). By contrast, of the 370 viruses originally isolated, nine agglutinated Sia $\alpha$ 2,6-TRBCs, albeit with different efficiencies (Supplementary Table 1). All nine viruses possessed mutations in the region targeted for random mutagenesis; one mutant also possessed an additional mutation (E119G) in an area that was not targeted for mutation. Most of the mutations clustered around the receptor-binding pocket (Fig. 1a). Several of the selected viruses possessed mutations known to increase binding to human-type receptors, including N186K (ref. 9), S227N (ref. 5) and Q226L (which confers human-type receptor binding together with G228S)<sup>15</sup> (all shown in blue in Fig. 1a). The identification of known determinants of human-type receptor-binding specificity from a library of random mutants validates our approach. Notably, our screen also identified mutations not previously associated with receptor-binding specificity.

Although viruses were diluted to  $\sim 0.5$  viruses per well for amplification in AX4 cells, we cannot exclude the possibility that some wells were infected with more than one virus, resulting in mixed populations. To confirm the significance of the identified mutations in HA for human-type receptor binding, the mutations were engineered into a VN1203/PR8 virus (possessing an avirulent HA cleavage site sequence, as described earlier). All nine mutants were generated; however, after two passages in MDCK cells, the S136N mutation reverted to the wild-type sequence. This mutant was excluded from further evaluation.

First, we confirmed the binding of the remaining eight variants to Sia $\alpha$ 2,6-TRBCs (Supplementary Table 1). For comparison, we included a VN1203/PR8 virus with two changes in its HA (Q226L and G228S) previously shown to have increased binding to Sia $\alpha$ 2,6Gal<sup>6,15</sup>. Indeed, compared to the wild-type VN1203/PR8 virus, the Q226L/G228S mutant displayed an increased ability to bind to

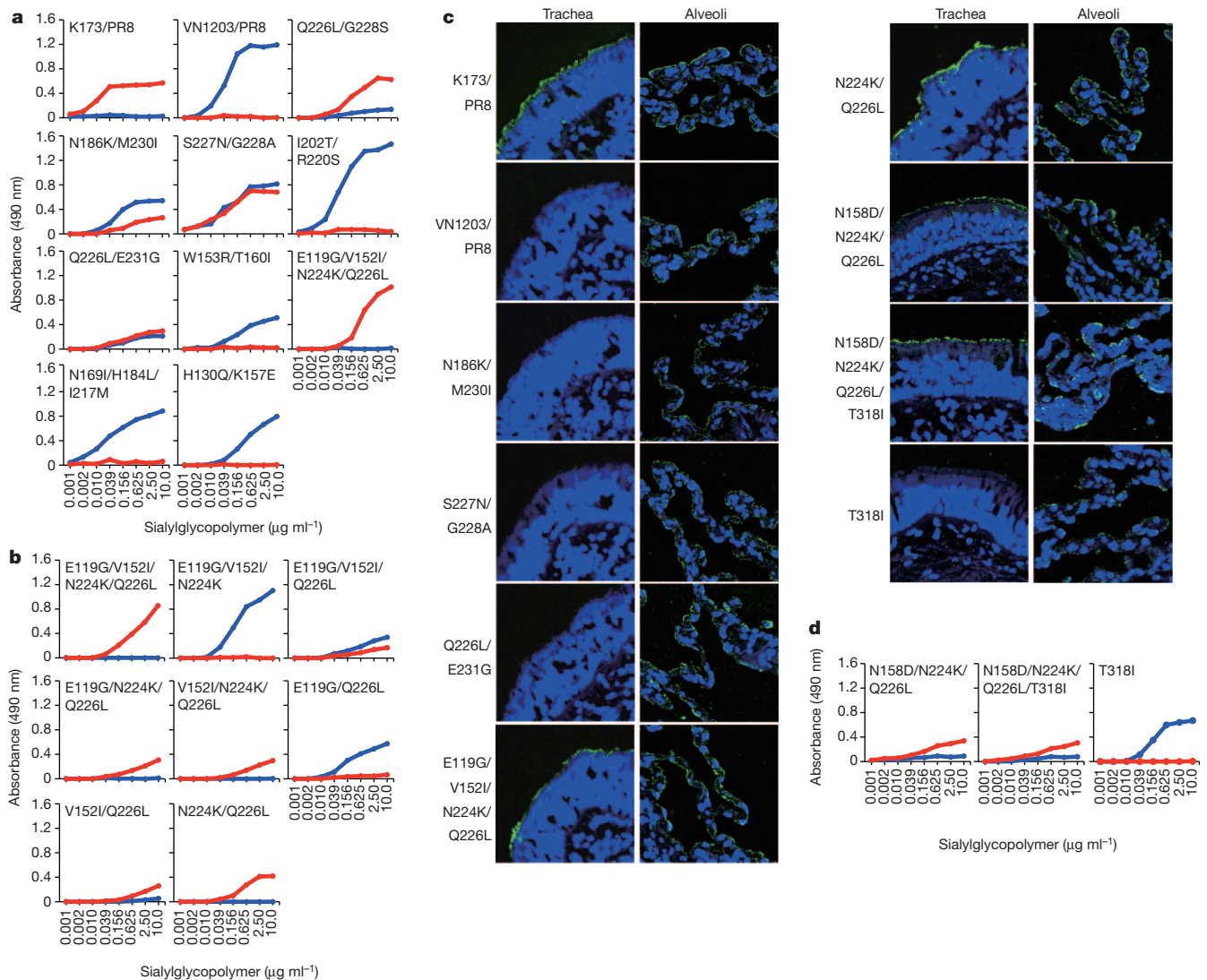
human-type receptors. For the recreated variants, haemagglutination titres were higher and slightly different from the initial characterization, which we attribute to biological differences (the initial characterization was carried out with non-concentrated cell culture supernatant and potentially mixed virus populations, whereas the recreated viruses were concentrated and purified) and to experimental differences (that is, differences between the TRBC batches or the efficiency of  $\alpha$ 2,3-sialidase treatment, or both). Collectively, however, these experiments demonstrate that this random mutagenesis approach allows the identification of hitherto unrecognized amino acid substitutions that permit avian virus HAs to bind to human-type receptors.

To characterize further the receptor-binding properties of the selected variants, we used solid-phase binding assays in which sialylglycopolymers were absorbed to plates, which were then incubated with virus (Fig. 2a). A virus possessing the HA and NA genes of the seasonal human A/Kawasaki/173/2001 (H1N1; K173) virus and the remaining genes from PR8 (K173/PR8) served as a control virus with typical human-type receptor specificity. Indeed, K173/PR8 preferentially bound to Sia $\alpha$ 2,6Gal. In contrast, VN1203/PR8 bound to only Sia $\alpha$ 2,3Gal. As reported elsewhere<sup>6,15</sup>, the Q226L/G228S mutations led to increased binding to Sia $\alpha$ 2,6Gal. Variants I202T/R220S, W153R/T160I, N169I/H184L/I217M and H130Q/K157E resembled VN1203/PR8 in their binding to glycans, despite the fact that these mutants weakly agglutinated Sia $\alpha$ 2,6-TRBCs (see Supplementary Table 1). These viruses may have bound to glycans on TRBCs that were different from Sia $\alpha$ 2,6Gal $\beta$ 1,4GlcNAc used in this study. However, variants N186K/M230I, S227N/G228A and Q226L/E231G showed an appreciable increase in binding to Sia $\alpha$ 2,6Gal but also retained binding capacity for Sia $\alpha$ 2,3Gal. Of all of the variants tested, only E119G/V152I/N224K/Q226L exhibited specificity for only Sia $\alpha$ 2,6Gal. Thus, only one H5 HA variant with receptor-binding capability akin to that of seasonal influenza viruses was isolated from the library screen of  $2.1 \times 10^6$  viruses. To identify the amino acid change(s) responsible for the conversion from Sia $\alpha$ 2,3Gal to Sia $\alpha$ 2,6Gal recognition in the E119G/V152I/N224K/Q226L virus HA, we tested the amino acid changes at positions 119, 152, 224 and 226 individually and in various combinations. Solid-phase binding assays demonstrated that the N224K/Q226L combination is critical for the shift from Sia $\alpha$ 2,3Gal to Sia $\alpha$ 2,6Gal recognition (Fig. 2b); Q226L in combination with V152I also conferred weak binding to  $\alpha$ 2,6-glycans.



**Figure 1 | Localization of amino acid changes identified in this study on the three-dimensional structure of the monomer of VN1203 HA (Protein Data Bank accession 2FK0)<sup>15</sup>.** **a**, Close-up view of the globular head of VN1203 HA. Mutations known to increase affinity to human-type receptors are shown in blue. Amino acid changes not previously known to affect receptor binding are shown in green. Additional mutations that occurred in the HA of H5 avian-

human reassortant viruses during replication and/or transmission in ferrets are shown in red. **b**, The positions of four mutations in the HA of H5 transmissible reassortant mutant virus, HA(N158D/N224K/Q226L/T318I)/CA04, are highlighted in red. The fusion peptide of HA is shown in cyan. All mutations are shown with H3 numbering. Images were created with MacPymol (<http://www.pymol.org/>).



**Figure 2 | Characterization of the receptor-binding properties of isolated viruses.** **a**, Binding of VN1203 mutants to sialylglycopolymers in solid-phase assays. A human virus (K173/PR8), an avian virus (VN1203/PR8) and mutant VN1203/PR8 viruses were compared for their ability to bind to sialylglycopolymers containing either  $\alpha$ 2,3-linked (blue) or  $\alpha$ 2,6-linked (red) sialic acids. **b**, Identification of mutations that confer binding to human-type receptors. **c**, Binding of VN1203 mutant viruses to human respiratory tissues. K173/PR8, VN1203/PR8 and mutant VN1203/PR8 viruses were incubated

To assess the effect of enhanced  $\alpha$ 2,6-glycan recognition on the attachment of viruses to human respiratory tracts, sections of tracheal and lung tissues were exposed to K173/PR8 (human-type receptor binder), VN1203/PR8 (avian-type receptor binder) and mutant VN1203/PR8 viruses (Fig. 2c). Because the N186K/M230I, S227N/G228A, Q226L/E231G, E119G/V152I/N224K/Q226L and N224K/Q226L mutants exhibited appreciable binding to Sia $\alpha$ 2,6Gal (Fig. 2a, b), the attachment of these mutants was also tested. On tracheal sections, the K173/PR8 virus bound extensively to ciliated epithelial cells (Fig. 2c and Supplementary Fig. 3), whereas the VN1203/PR8 virus bound poorly. By contrast, on lung sections, both viruses bound extensively to the alveolar epithelial surface (both type I and II pneumocytes; Fig. 2c and Supplementary Fig. 4). The binding patterns of these viruses correlate with the distribution of Sia $\alpha$ 2,3Gal (that is, avian-type receptors; present in lung epithelia) and Sia $\alpha$ 2,6Gal (that is, human-type receptors; present in both trachea and lung epithelia) on the tissues, as observed with lectin staining<sup>18</sup> (Supplementary Fig. 5). Like the human K173/PR8 virus, the E119G/V152I/N224K/Q226L and N224K/Q226L

mutants exhibited strong binding to the ciliated epithelial cells of the trachea (Fig. 2c and Supplementary Fig. 3). By contrast, the N186K/M230I, S227N/G228A and Q226L/E231G mutants displayed little-to-no binding to tracheal epithelia (Fig. 2c), despite their binding to Sia $\alpha$ 2,6Gal (Fig. 2a). A number of sialylated oligosaccharides with differing branching patterns and chain lengths are thought to be present on the cell surface<sup>19</sup>. We therefore speculate that the mutants can recognize a short glycan structure such as Sia $\alpha$ 2,6Gal $\beta$ 1,4GlcNAc, but may not recognize longer, more complex glycan structures, which are possibly required for binding to human tracheal epithelium. On the other hand, all mutants bound to alveolar epithelial cells (both type I and II pneumocytes; Fig. 2c and Supplementary Fig. 4). When the tissue sections were pre-treated with *Arthrobacter ureafaciens* sialidase (which cleaves all non-reducing terminally branched and unbranched sialic acids), virus binding to the tissues was substantially reduced (Supplementary Fig. 6a–c), confirming the sialic acid binding specificity of the virus. These data indicate that alterations in the receptor specificity of the E119G/V152I/N224K/Q226L and N224K/Q226L

mutants exhibited strong binding to the ciliated epithelial cells of the trachea (Fig. 2c and Supplementary Fig. 3). By contrast, the N186K/M230I, S227N/G228A and Q226L/E231G mutants displayed little-to-no binding to tracheal epithelia (Fig. 2c), despite their binding to Sia $\alpha$ 2,6Gal (Fig. 2a). A number of sialylated oligosaccharides with differing branching patterns and chain lengths are thought to be present on the cell surface<sup>19</sup>. We therefore speculate that the mutants can recognize a short glycan structure such as Sia $\alpha$ 2,6Gal $\beta$ 1,4GlcNAc, but may not recognize longer, more complex glycan structures, which are possibly required for binding to human tracheal epithelium. On the other hand, all mutants bound to alveolar epithelial cells (both type I and II pneumocytes; Fig. 2c and Supplementary Fig. 4). When the tissue sections were pre-treated with *Arthrobacter ureafaciens* sialidase (which cleaves all non-reducing terminally branched and unbranched sialic acids), virus binding to the tissues was substantially reduced (Supplementary Fig. 6a–c), confirming the sialic acid binding specificity of the virus. These data indicate that alterations in the receptor specificity of the E119G/V152I/N224K/Q226L and N224K/Q226L

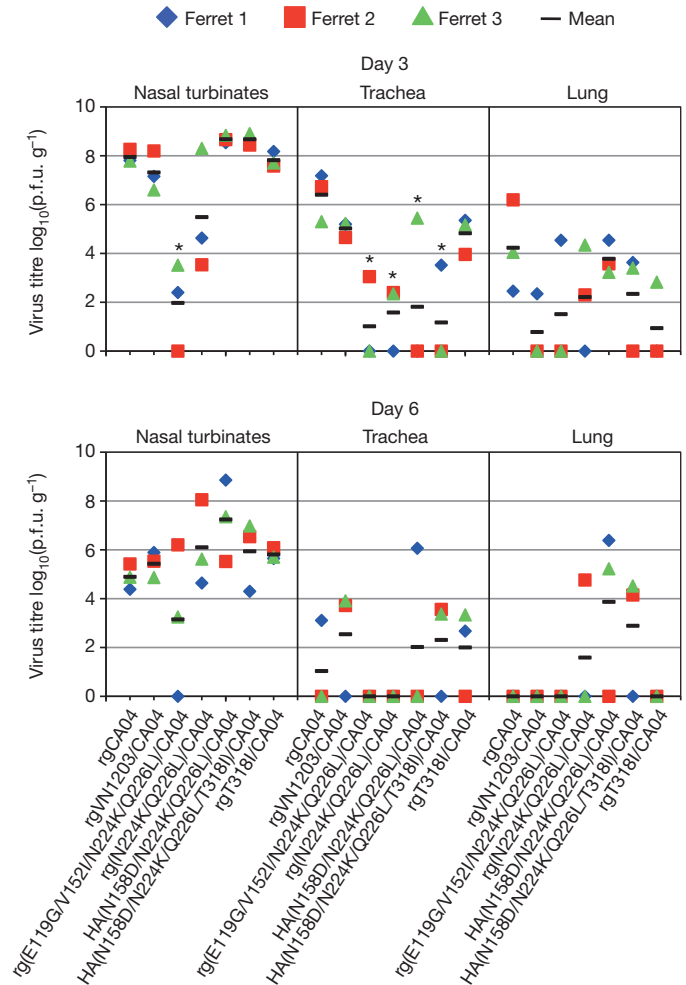


mutants have profound effects on virus attachment to human respiratory epithelium.

In an avian H3 HA, the Q226L mutation changed the binding preference from avian- to human-type<sup>20</sup>. A previous study found that the Q226L mutation on an H5 HA does not confer efficient binding to  $\alpha$ 2,6-glycans in a glycan array<sup>15</sup>; however, when tested in combination with G228S, increased binding to human-type receptors, but not a complete switch from avian- to human-type receptor-binding specificity, was observed<sup>15</sup>. By contrast, here we found that Q226L in combination with N224K resulted in a switch from Sia $\alpha$ 2,3Gal to Sia $\alpha$ 2,6Gal binding in an H5 HA and allowed virus binding to human tracheal epithelia (Fig. 2c). The receptor-binding domain of HA is formed by the 190-helix at the top of HA, the 220-loop at the edge of the globular head, and the 130-loop at the other edge of the globular head (Fig. 1a). Crystal structure analysis revealed that the 220-loop of avian H5 HA is closer to the opposing 130-loop than in human H3 HA, indicating that a wider binding site for human H3 HA, compared to that of avian H5 HA, may be required to optimize contacts with the larger Sia $\alpha$ 2,6-glycans<sup>21</sup>. N224 lies on the turn leading into the 220-loop, adjacent to position 226 (Fig. 1a). Replacement of N224 may alter the orientation of the 220-loop and thus optimize contacts between L226 and Sia $\alpha$ 2,6Gal-containing receptors, thereby increasing the preference for  $\alpha$ 2,6 linkages.

Recent studies reported that 2009 pandemic H1N1 and H5N1 viruses show high genetic compatibility<sup>22,23</sup>. These two viruses have been isolated from pigs<sup>24–28</sup>, which have been considered as ‘mixing vessels’ for the reassortment of avian, swine and human strains. Thus, the coexistence of H5N1 and 2009 pandemic H1N1 viruses could provide an opportunity for the generation of transmissible H5 avian-human reassortants in mammals. Therefore, we generated reassortant viruses possessing the mutant VN1203 HAs generated above, and the seven remaining gene segments from a prototype 2009 pandemic H1N1 virus (A/California/04/2009, CA04). Experiments with viruses possessing the wild-type HA cleavage site were performed in enhanced biosafety level 3 (BSL3+) containment laboratories approved for such use by the Centers for Disease Control and Prevention (CDC) and the United States Department of Agriculture (USDA). Because efficient human-to-human transmission is a critical feature of pandemic influenza viruses, we examined the growth and transmissibility of reassortant viruses in ferrets, which are widely accepted as an animal model for influenza virus transmissibility and pathogenesis studies. Because the E119G/V152I/N224K/Q226L and N224K/Q226L variants bound extensively to human tracheal epithelia (Fig. 2c), we generated by reverse genetics (rg) three H5 reassortant viruses possessing the VN1203 HA or mutant HAs (all with the wild-type multibasic cleavage site) and the remaining genes from the CA04 virus. The VN1203 HA mutants tested included the one containing four mutations, E119G, V152I, N224K and Q226L (designated rg(E119G/V152I/N224K/Q226L)/CA04), and another containing two mutations, N224K and Q226L (designated rg(N224K/Q226L)/CA04).

To determine whether the introduced HA mutations affected the replication of the H5 reassortant viruses, six ferrets were inoculated intranasally with  $10^6$  plaque-forming units (p.f.u.) of virus. On day 3 after infection, a recombinant virus whose genes all came from CA04, rgCA04, replicated efficiently in the respiratory organs of infected animals, and was isolated from the colon, but not from any other organs tested (Fig. 3 and Supplementary Table 2). A virus possessing H5 VN1203 HA and the remaining genes from CA04 (designated rgVN1203/CA04) replicated to titres comparable to those of rgCA04 in nasal turbinates, but substantially less in the lungs. By contrast, the two H5 reassortant viruses with HA mutations (rg(E119G/V152I/N224K/Q226L)/CA04 and rg(N224K/Q226L)/CA04) were severely limited in their replicative ability in trachea. Although virus titres in nasal turbinates and lung were not statistically different between rg(N224K/Q226L)/CA04 and rgCA04, the virus titre in nasal turbinates was significantly lower in animals inoculated with rg(E119G/



**Figure 3 | Virus replication in respiratory organs.** Ferrets were infected intranasally with  $10^6$  p.f.u. of virus. Three ferrets per group were killed on days 3 and 6 after infection for virus titration. Virus titres in nasal turbinates, trachea and lung were determined by use of a plaque assay on MDCK cells. Horizontal bars show the mean. Asterisks indicate virus titres significantly different from that of rgCA04 (Dunnett's test;  $P < 0.05$ ).

V152I/N224K/Q226L)/CA04 than in animals inoculated with rgCA04 (Dunnett's test;  $P = 0.0002$ ; Fig. 3). Notably, rgVN1203/CA04 (avian-type receptor binder) replicated efficiently in nasal turbinates of ferrets, which have a similar sialic acid receptor distribution pattern to that of the human respiratory tract<sup>29,30</sup>. The reason for this discrepancy is unclear; however, replication of avian H5N1 viruses in ferret nasal turbinates has been reported<sup>12,13</sup>.

Although virus titres in respiratory organs were generally lower on day 6 after infection than on day 3 after infection, rg(N224K/Q226L)/CA04 still showed high levels of replication at day 6 after infection; titres in nasal turbinates ranged from  $10^{4.6}$  to  $10^{8.1}$  p.f.u. g<sup>-1</sup> (Fig. 3). Sequence analysis of viruses in nasal turbinates on day 6 after infection revealed that viruses in ferret 2 and ferret 3 possessed N158D and N158K mutations in their HA (in addition to the original two mutations), respectively, leading to the loss of the glycosylation site at position 158 (that is, 158N-S-T to 158D-S-T or 158K-S-T; Fig. 1a and Supplementary Table 3). In nasal turbinates on day 6 after infection, the titre of the virus with the N158D/N224K/Q226L mutations ( $10^{8.1}$  p.f.u. g<sup>-1</sup>; see Fig. 3, ferret 2 of rg(N224K/Q226L)/CA04) was approximately four orders of magnitude higher than that of the original rg(N224K/Q226L)/CA04 ( $10^{4.6}$  p.f.u. g<sup>-1</sup>; Fig. 3, ferret 1 of rg(N224K/Q226L)/CA04), whereas the virus with the N158K/N224K/Q226L mutations ( $10^{5.6}$  p.f.u. g<sup>-1</sup>; Fig. 3, ferret 3 of rg(N224K/Q226L)/CA04) grew to one order of magnitude higher than

the original mutant. These data indicate that the additional mutation N158D improved the replication of rg(N224K/Q226L)/CA04 in ferrets. To test the effect of this mutation on the replication of H5 reassortant viruses in ferrets, we examined the replicative ability of a virus with the triple N158D/N224K/Q226L HA substitutions in ferrets. This HA(N158D/N224K/Q226L)/CA04 virus replicated efficiently in infected animals, except in the trachea (Fig. 3 and Supplementary Table 2). On day 3 after infection, this virus was isolated from the brain of two of the three animals tested, although we did not observe neurological signs in these animals. These results indicate that the N158D mutation contributed to the efficient growth in the nasal turbinates of ferrets of an H5 reassortant virus with the N224K/Q226L mutations. Removal of the glycosylation site at position 158 has been reported to result in enhanced binding of H5N1 viruses to human-type receptors in combination with the Q226L/G228S mutations<sup>7</sup>. A previous study showed that H5N1 viruses lacking this glycosylation site transmit efficiently by direct contact among guinea-pigs<sup>31</sup>. By contrast, H5N1 viruses that acquire this glycosylation site lose the ability to transmit among guinea-pigs. Therefore, we speculated that the loss of the glycosylation site in HA(N158D/N224K/Q226L)/CA04 virus may affect its transmissibility in ferrets.

To assess the ability of H5 reassortant viruses with human-type receptor specificity to transmit between ferrets, we placed naive ferrets in wireframe cages next to ferrets inoculated with  $10^6$  p.f.u. of rgCA04, rgVN1203/CA04, rg(N224K/Q226L)/CA04, or HA(N158D/N224K/Q226L)/CA04 (Supplementary Fig. 7). Similar to previous experiments<sup>32</sup>, rgCA04 was efficiently transmitted via respiratory droplets to all three contact ferrets, as evidenced by the detection of virus in nasal washes and haemagglutination inhibition (HI) antibody in these animals (Table 1 and Fig. 4). By contrast, rgVN1203/CA04 and rg(N224K/Q226L)/CA04 were not transmitted; neither virus shedding nor seroconversion was detected in any contact animals, despite the binding of the latter to Sia $\alpha$ 2,6Gal. This result was consistent with that of previous studies in which human-type receptor recognition was shown to be necessary but not sufficient for respiratory droplet transmission of an H5N1 virus in a ferret model<sup>12,14</sup>. In the HA(N158D/N224K/Q226L)/CA04-inoculated group, virus was recovered from two of the six contact ferrets (pairs 1 and 2) between days 5 and 7 after contact. Moreover, seroconversion was detected in five animals including those from which virus was recovered. No animals died in the course of these transmission experiments. This finding demonstrates the generation of an H5 HA that supports virus transmission by respiratory droplets among ferrets.

To determine whether additional mutations occurred in the HA of HA(N158D/N224K/Q226L)/CA04 during transmission, viral RNA was analysed from nasal washes of inoculated and contact ferrets (Fig. 4 and Supplementary Table 4). On day 5 after infection, the A242S and T318I mutations in HA were present in five (pairs 1, 3, 4, 5 and 6) and one (pair 2) of the six inoculated animals, respectively. Viruses derived from the contact animals of pair 1 on day 7 after contact had two changes in HA

(K193N and A242S) (Fig. 1a), whereas those derived from the contact animals of pair 2 contained a single change in HA (T318I) (Fig. 1b), indicating that additional changes in HA occurred during the infection of ferrets with HA(N158D/N224K/Q226L)/CA04. No mutations in the remaining genes were detected in any of these viruses from nasal washes compared with the CA04 virus sequences.

Because HA(N158D/N224K/Q226L)/CA04 was isolated from only one-third of the contact animals, we isolated a virus from the nasal wash of the contact ferret that shed a high titre ( $10^{7.5}$  p.f.u. ml<sup>-1</sup>) of virus on day 7 after contact (pair 2) (Fig. 4d) to evaluate the replication and transmissibility of that virus in ferrets. This mutant virus, designated HA(N158D/N224K/Q226L/T318I)/CA04, replicated efficiently in the nasal turbinates and was isolated from brain tissue (Fig. 3 and Supplementary Table 2). In the transmission study, four of the six contact ferrets were positive for virus between days 3 and 7 after contact, and all contact animals were seropositive; no animals died in the course of the transmission experiments (Table 1; Fig. 4e and Supplementary Fig. 8). Notably, this transmission pattern is comparable to that of the 1918 pandemic H1N1 virus when tested under the same experimental conditions; the 1918 pandemic virus was recovered from the nasal wash of two of three contact animals (our own unpublished data). Sequence comparison of viruses from inoculated and contact animals identified mutations at positions 225 and 242 as well as a reversion at position 224 (Fig. 1a and Supplementary Table 5) (in addition to the original four mutations) although the 224 reversion was found only in viruses from inoculated ferrets. Collectively, these findings demonstrate that four amino acid substitutions (N158D/N224K/Q226L/T318I) in H5 HA confer efficient respiratory droplet transmission in ferrets to a virus possessing an H5 HA in a 2009 pandemic H1N1 backbone. We also confirmed that recombinant viruses possessing the three HA mutations N158D, N224K and Q226L, or the four HA mutations N158D, N224K, Q226L and T318I, and the NA of VN1203 in a PR8 background (designated N158D/N224K/Q226L or N158D/N224K/Q226L/T318I, respectively) preferentially bind to Sia $\alpha$ 2,6Gal and attach to human tracheal epithelia (Fig. 2c, d).

HA(N158D/N224K/Q226L/T318I)/CA04 transmitted by respiratory droplet more efficiently than HA(N158D/N224K/Q226L)/CA04, raising the possibility that the T318I mutation is involved in the efficient transmission of avian H5N1/pandemic H1N1 reassortants. To explore the functional role of this mutation in respiratory droplet transmission, we generated an H5 reassortant expressing the H5 HA with the T318I mutation and examined its receptor-binding specificity and transmissibility. This reassortant (designated rgT318I/CA04) bound to only Sia $\alpha$ 2,3Gal and showed little binding to human tracheal epithelia (Fig. 2c, d). rgT318I/CA04 did not transmit via respiratory droplet among ferrets (Table 1 and Fig. 4f), although it replicated in nasal turbinates and trachea as efficiently as rgCA04 (Fig. 3 and Supplementary Table 2). These results indicate that the T318I mutation alone is not sufficient for H5 reassortant viruses to transmit efficiently among ferrets.

**Table 1 | Transmission in ferrets inoculated with H5 avian-human reassortant viruses**

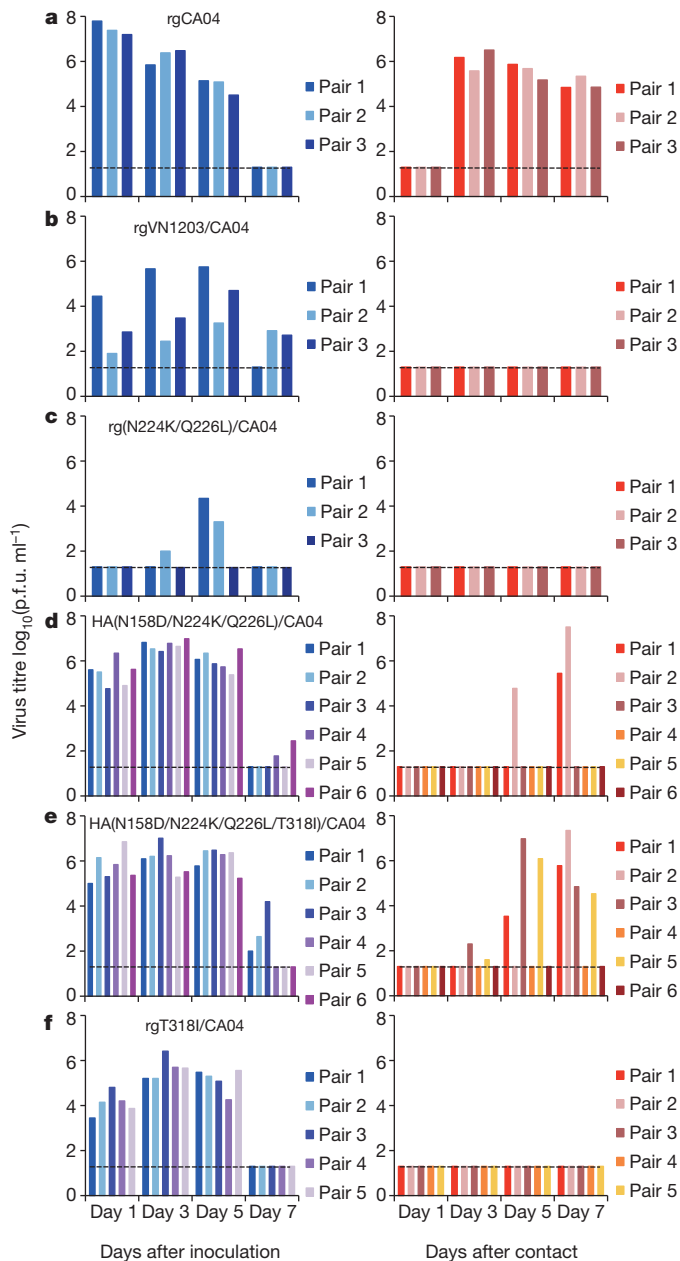
Virus	Inoculated ferrets			Contact ferrets	
	Weight loss (%) <sup>*</sup>	Peak virus titre in nasal wash (mean log <sub>10</sub> (p.f.u. ml <sup>-1</sup> )) (days after inoculation)	Seroconversion (positive and total numbers) (HI titre) <sup>†</sup>	Virus detection in nasal wash (positive and total numbers)	Seroconversion (positive and total numbers) (HI titre)
rgCA04	3 of 3 (15.1)	7.5 (1)	3 of 3 ( $\geq 1,280$ , $\geq 1,280$ , $\geq 1,280$ )	3 of 3	3 of 3 ( $\geq 1,280$ , $\geq 1,280$ , $\geq 1,280$ )
rgVN1203/CA04	3 of 3 (5.9)	5.3 (5)	3 of 3 (80, 40, 80)	0 of 3	0 of 3 (<10, <10, <10)
rg(N224K/Q226L)/CA04	2 of 3 (7.8) <sup>‡</sup>	3.9 (5)	3 of 3 ( $\geq 1,280$ , $\geq 1,280$ , $\geq 1,280$ )	0 of 3	0 of 3 (<10, <10, <10)
HA(N158D/N224K/Q226L)/CA04	6 of 6 (5.7)	6.7 (3)	6 of 6 (640, $\geq 1,280$ , $\geq 1,280$ , 640, $\geq 1,280$ , $\geq 1,280$ )	2 of 6	5 of 6 (160, 320, 20, 160, 40, <10)
HA(N158D/N224K/Q226L/T318I)/CA04	6 of 6 (9.8)	6.1 (5)	6 of 6 ( $\geq 1,280$ , $\geq 1,280$ , 640, $\geq 1,280$ , $\geq 1,280$ , $\geq 1,280$ )	4 of 6	6 of 6 (640, 640, $\geq 1,280$ , 80, $\geq 1,280$ , 320)
rgT318I/CA04	3 of 5 (1.5) <sup>§</sup>	5.6 (3)	5 of 5 (40, 20, 20, 40, 40)	0 of 5	0 of 5 (<10, <10, <10, <10, <10)

<sup>\*</sup> Maximum percentage weight loss is shown.

<sup>†</sup> Haemagglutination inhibition (HI) assays were carried out with homologous virus and turkey red blood cells.

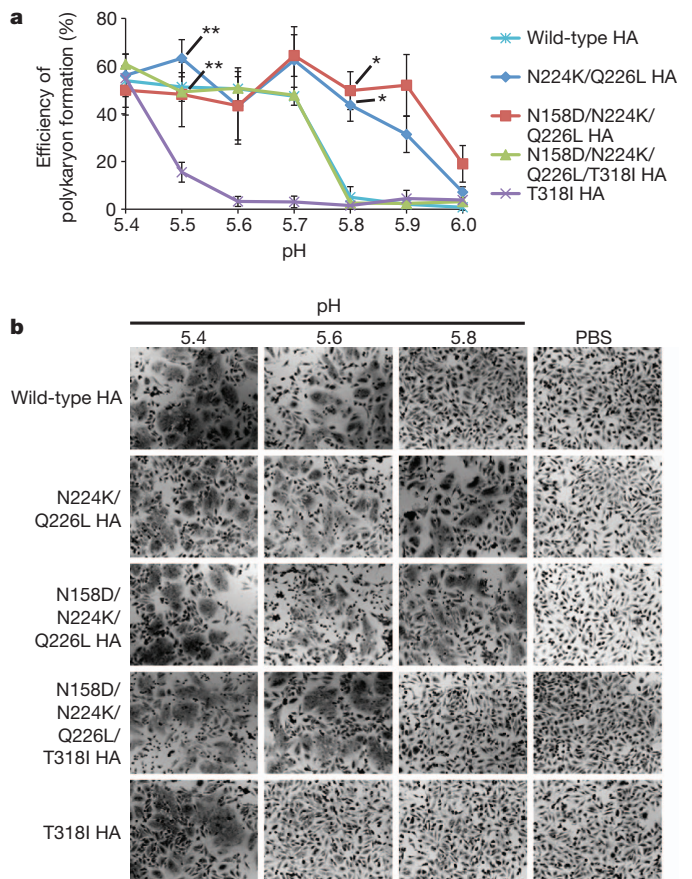
<sup>‡</sup> One animal did not lose any body weight.

<sup>§</sup> Two animals did not lose any body weight.



**Figure 4 | Respiratory droplet transmission of H5 avian-human reassortant viruses in ferrets.** a–f, Groups of three, five, or six ferrets were inoculated intranasally with 10<sup>6</sup> p.f.u. of rgCA04 (a), rgVN1203/CA04 (b), rg(N224K/Q226L)/CA04 (c), HA(N158D/N224K/Q226L)/CA04 (d), HA(N158D/N224K/Q226L/T318I)/CA04 (e), or rgT318I/CA04 (f). One day after infection, three, five, or six naive ferrets were placed in adjacent cages. Nasal washes were collected every other day from both inoculated (left panel) and contact (right panel) animals for virus titration. Virus titres in organs were determined by using a plaque assay on MDCK cells. The lower limit of detection is indicated by the horizontal dashed line.

Influenza virus HA protein has membrane-fusion as well as receptor-binding activity. Notably, in the three-dimensional model of influenza A virus HA, residue 318 is located proximally to the fusion peptide (Fig. 1b), which has key roles in the membrane fusion process. To assess the effect of HA mutations on low-pH-induced membrane fusion activity, we examined the pH at which the fusion activity of wild-type and mutant HA was activated (Fig. 5). The wild-type HA had a threshold for membrane fusion of pH 5.7; the N224K/Q226L and N158D/N224K/Q226L mutations raised the threshold for fusion to >pH 5.9, whereas the T318I mutation reduced the threshold for fusion to pH 5.5. The N158D/N224K/Q226L/T318I mutations showed

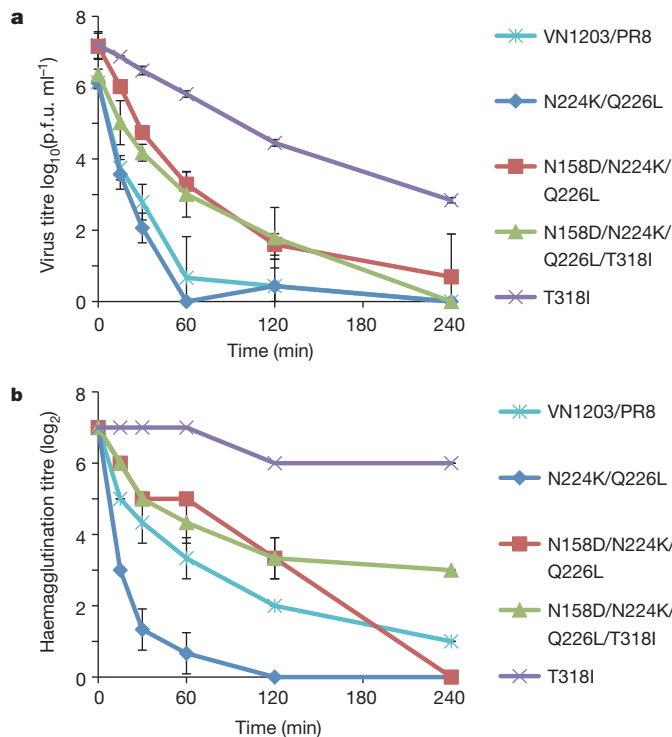


**Figure 5 | Polykaryon formation by HeLa cells expressing wild-type or mutant HAs after acidification at low pH.** a, The efficiency of polykaryon formation over a pH range of 5.4–6.0 was estimated from the number of nuclei in polykaryons divided by the total number of nuclei in the same field. The mean and standard deviations determined from five randomly chosen fields of cell culture are shown. Single asterisks indicate values significantly different between the wild-type HA and the N224K/Q226L or N158D/N224K/Q226L HA (Tukey test;  $P < 0.05$ ). The double asterisk indicates values significantly different between the T318I HA and the N224K/Q226L or N158D/N224K/Q226L HA (Tukey test;  $P < 0.05$ ). b, Representative fields of cells expressing the indicated HAs and exposed to pH 5.4, 5.6, or 5.8 are shown. Images were taken at  $\times 10$  magnification.

wild-type fusogenic properties (that is, a threshold at pH 5.7). The HA of influenza virus undergoes a low-pH-dependent conformational change, which is required for fusion of the viral envelope with the target membrane<sup>33</sup>. Such a conformational change to a fusion-active form can also lead to viral inactivation. Therefore, sustained and efficient human-to-human transmission of virus may require a certain level of stability of the HA protein in an acidic environment, as the pH of human nasal mucosa, where human influenza viruses replicate primarily, is approximately pH 5.5–6.5 (ref. 34). Our findings suggest that an increase in the pH threshold for fusion as a result of the N224K/Q226L mutations that shift the HA receptor recognition from avian-type to human-type may reduce HA protein stability; however, the T318I mutation decreases the pH threshold for fusion activity, resulting in a stable mutant HA.

Because heat treatment at neutral pH is also known to promote a fusogenic form of HA protein<sup>35,36</sup> and serve as a surrogate assay for HA stability<sup>37</sup>, we next tested whether the HA mutations described above affect the heat stability of the HA protein. Wild-type and mutant HA viruses were incubated at 50 °C for various times, after which the loss of infectivity and haemagglutination activity were determined. The wild-type and N224K/Q226L viruses lost most of their infectivity by heating for 60 min ( $>5.5$ -log<sub>10</sub> decrease in titre; Fig. 6a), whereas the

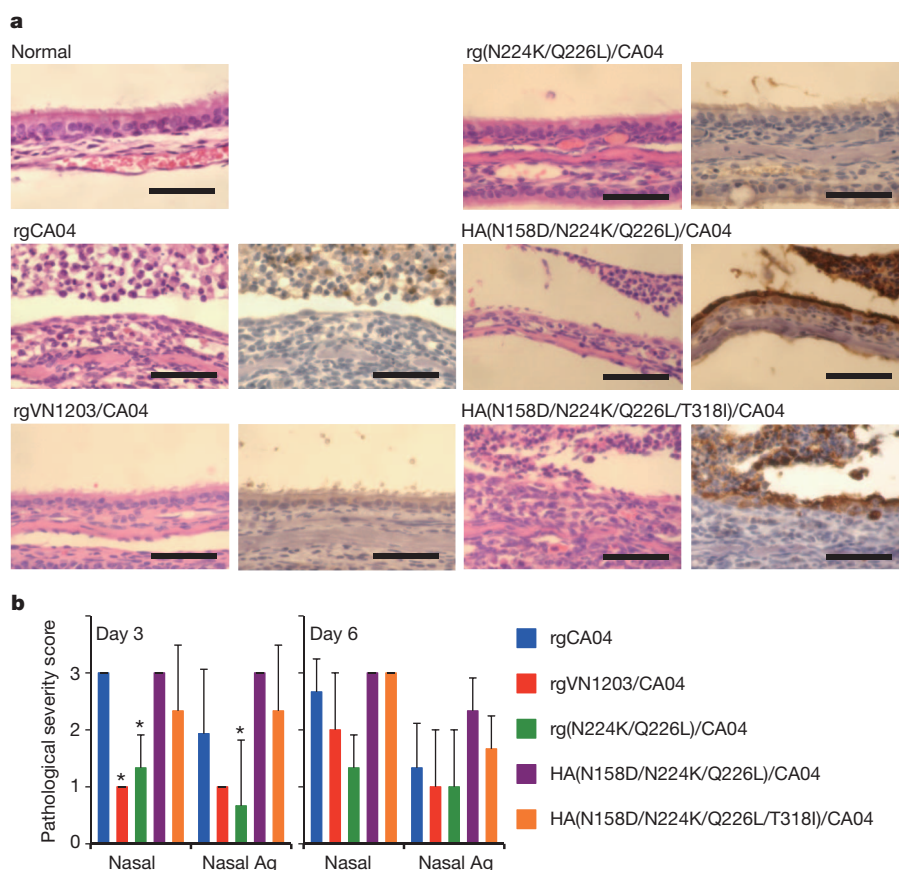




**Figure 6 | Effect of heat treatment on the infectivity and haemagglutination activity of viruses.** Aliquots of a virus stock containing 128 HA units were incubated for the times indicated at 50 °C. **a**, Virus titres in heat-treated samples were determined by plaque assays on MDCK cells. **b**, Haemagglutination titres in heat-treated samples were determined by using haemagglutination assays with 0.5% TRBCs. Each point represents the mean  $\pm$  standard deviation from triplicate experiments.

N158D/N224K/Q226L and N158D/N224K/Q226L/T318I mutants exhibited considerable tolerance to high temperature (3.9- and 3.4- $\log_{10}$  decrease after a 60-min incubation, respectively) and the T318I mutant was most resistant (only a 1.4- $\log_{10}$  decrease under the same conditions). In haemagglutination assays, the N224K/Q226L mutant HA lost activity more rapidly than did the wild-type HA, and N158D/N224K/Q226L lost activity more rapidly than did the N158D/N224K/Q226L/T318I mutant (Fig. 6b). Thus, addition of the N158D mutation to the N224K/Q226L HA increased HA stability and subsequent addition of the fourth mutation, T318I, rendered the HA protein even more stable. Taken together, these results suggest that the addition of the T318I mutation to H5 HAs that preferentially recognize human-type receptors restores HA protein stability, thereby allowing a virus carrying the N158D/N224K/Q226L/T318I mutations in HA to transmit efficiently via respiratory droplet among ferrets. In conclusion, a fine balance of mutations affecting different functions in HA (such as receptor-binding specificity and HA stability) may be critical to confer transmissibility in ferrets.

We next compared the pathogenicity in ferrets of H5 avian-human reassortants with that of the pandemic H1N1 virus CA04 (Fig. 7, Supplementary Information and Supplementary Figs 9–11). The control virus, rgCA04, caused substantial body weight loss (15.1%) (Table 1 and Supplementary Fig. 9). By contrast, the four reassortant viruses caused only modest weight loss (<10%) in most of the animals. However, no statistically significant differences in body weight loss were found between the reassortant viruses and rgCA04. Pathological examination revealed similar histological changes and levels of viral antigens in the nasal mucosa of rgCA04-, HA(N158D/N224K/Q226L)/CA04- and HA(N158D/N224K/Q226L/T318I)/CA04-infected ferrets (Fig. 7a, b). In the rgVN1203/CA04 and rg(N224K/Q226L)/CA04 groups, however, less tissue damage was found in the nasal mucosa compared with the rgCA04 group on day 3 after infection (Dunnett's test;  $P = 0.0057$  and  $0.0175$ , respectively; Fig. 7b). In addition, all three



**Figure 7 | Pathological analyses of H5 avian-human reassortant viruses.** **a**, Representative histological changes in nasal turbinates from influenza-virus-infected ferrets. Three ferrets per group were infected intranasally with  $10^6$  p.f.u. of virus, and tissues were collected on day 3 after infection for pathological examination. Uninfected ferret tissues served as negative controls (normal). Left panel, haematoxylin-and-eosin staining. Right panel, immunohistochemical staining for viral antigen detection (brown staining). Scale bars, 50  $\mu\text{m}$ . **b**, Pathological severity scores in infected ferrets. To represent comprehensive histological changes, respiratory tissue slides were evaluated by scoring the pathological changes and viral antigen expression levels. The pathological scores were determined for each animal in each group ( $n = 3$  per group on days 3 and 6 after infection) using the following scoring system: 0, no pathological change/antigen negative; 1, affected area (<30%) or only interstitial lesion/rare viral antigens; 2, affected area (<80%,  $\geq 30\%$ )/moderate viral antigens; 3, severe lesion ( $\geq 80\%$ )/many viral antigens. Nasal, pathological changes in the nasal mucosa; nasal Ag, viral antigens in the nasal mucosa. Asterisks indicate virus pathological scores significantly different from that of rgCA04 (Dunnett's test;  $P < 0.05$ ). Error bars denote standard deviation.

viruses caused lung lesions (Supplementary Information and Supplementary Figs 10 and 11).

To assess whether current control measures may be effective against the H5 transmissible reassortant mutant virus, we examined the reactivity of sera from individuals vaccinated with an H5N1 prototype vaccine<sup>38</sup> against a virus possessing the N158D/N224K/Q226L/T318I mutations in HA. We found that pooled human sera from individuals immunized with this vaccine reacted with the virus possessing the mutant H5 HA (N158D/N224K/Q226L/T318I) at a higher titre than with a wild-type H5 HA virus (VN1203/PR8; Supplementary Table 6), indicating that current H5N1 vaccines would be efficacious against the H5 transmissible reassortant mutant virus. In addition, the H5 transmissible reassortant mutant virus (HA(N158D/N224K/Q226L/T318I)/CA04) was highly susceptible to a licensed NA inhibitor, oseltamivir (Supplementary Table 7). These experiments show that appropriate control measures would be available to combat the transmissible virus described in this study.

Currently, we do not know whether the mutations that we identified in this study that allowed the HA(N158D/N224K/Q226L/T318I)/CA04 virus to be transmissible in ferrets would also support sustained human-to-human transmission. In particular, we wish to emphasize that the transmissible HA(N158D/N224K/Q226L/T318I)/CA04 virus possesses seven segments (all but the HA segment) from a human pandemic 2009 H1N1 virus. Human-virus-characteristic amino acids in these seven segments may have critically contributed to the respiratory droplet transmission of the HA(N158D/N224K/Q226L/T318I)/CA04 virus in ferrets. Examples include amino acids in the PB2 polymerase protein that confer efficient replication in mammalian, but not avian, cells<sup>39–43</sup>. As the PB2 gene of the HA(N158D/N224K/Q226L/T318I)/CA04 virus is of human virus origin, the virus possesses high replicative ability in mammalian cells. In contrast, most avian virus PB2 proteins lack these human-type amino acids, although one of these changes (a glutamic-acid-to-lysine mutation at position 627) is found in highly pathogenic avian H5N1 viruses circulating in the Middle East<sup>44</sup>. As a second example, the viral NA gene may contribute to viral transmissibility. The NA protein cleaves  $\alpha$ -ketosidic linkages between a terminal sialic acid and an adjacent sugar residue, an activity that balances the sialic-acid-binding activity of HA. A recent study found that a human virus NA gene was critical to confer limited transmissibility to a mutant H5 avian-human reassortant virus<sup>14</sup>. In general, a human-type receptor recognizing H5 HA alone may not be sufficient to confer transmissibility in mammals, but may have to act together with other human-virus-characteristic traits (in PB2, NA, and/or other viral proteins). Therefore, at this point we cannot predict whether the four mutations in the H5 HA identified here would render a wholly avian H5N1 virus transmissible.

Three of the residues identified here (N224, Q226 and T318) have been strictly conserved among H5 HA proteins isolated since 2003. However, as H5N1 viruses continue to evolve and infect people, receptor-binding variants of H5N1 viruses, including avian-human reassortant viruses as tested here, may emerge. One of the four mutations we identified in our transmissible virus, the N158D mutation, results in loss of a glycosylation site. Many H5N1 viruses isolated in the Middle East, Africa, Asia and Europe do not have this glycosylation site. Therefore, only three nucleotide changes are needed for the HA of these viruses to support efficient transmission in ferrets. In addition, the H5N1 viruses circulating in these geographic areas also possess a glutamic-acid-to-lysine mutation at position 627 in the PB2 protein, which promotes viral replication in certain mammals, including humans<sup>40,45</sup>. Therefore, these viruses may be several steps closer to those capable of efficient transmission in humans and are of concern.

Our study highlights the pandemic potential of viruses possessing an H5 HA. Although current vaccines may protect against a virus similar to that tested here, the continued evolution of H5N1 viruses reinforces the need to prepare and update candidate vaccines to H5 viruses. The amino acid changes identified here will help individuals

conducting surveillance in regions with circulating H5N1 viruses (for example, Egypt, Indonesia, Vietnam) to recognize key residues that predict the pandemic potential of isolates. Rapid responses in a potential pandemic situation are essential in order to generate appropriate vaccines and initiate other public health measures to control infection. Furthermore, our findings are of critical importance to those making public health and policy decisions.

Our research answers a fundamental question in influenza research: can H5-HA-possessing viruses support transmission in mammals? Moreover, our findings have suggested that different mechanisms (that is, receptor-binding specificity and HA stability) may act in concert for efficient transmissibility in mammals. This knowledge will facilitate the identification of additional mutations that affect viral transmissibility; the monitoring of this expanded set of changes in natural isolates may improve our ability to assess the pandemic potential of H5N1 viruses. Thus, although a pandemic H5N1 virus may not possess the amino acid changes identified in our study, the findings described here will advance our understanding of the mechanisms and evolutionary pathways that contribute to avian influenza virus transmission in mammals.

## METHODS SUMMARY

**Viruses.** All recombinant viruses were generated by using reverse genetics essentially as described previously<sup>16</sup>. All experiments with the viruses possessing the wild-type HA cleavage site were performed in an enhanced biosafety level 3 (BSL3+) containment laboratory approved for such use by the CDC and the USDA.

**Infection and transmission in ferrets.** Six-ten-month-old female ferrets (Triple F Farms) were intramuscularly anaesthetized and intranasally inoculated with  $10^6$  p.f.u. (500  $\mu$ l) of virus. On days 3 and 6 after infection, ferrets were killed for virological and pathological examinations. The virus titres in various organs were determined by use of plaque assays in MDCK cells.

For transmission studies in ferrets, animals were housed in adjacent transmission cages that prevented direct and indirect contact between animals but allowed spread of influenza virus through the air (Showa Science; Supplementary Fig. 7). Ferrets were intranasally inoculated with  $10^6$  p.f.u. (500  $\mu$ l) of virus (inoculated ferrets). Twenty-four hours after infection, naive ferrets were each placed in a cage adjacent to an inoculated ferret (contact ferrets). To assess viral replication in the nasal turbinates, we determined viral titres in nasal washes collected from virus-inoculated and contact ferrets on day 1 after inoculation or co-housing, respectively, and then every other day. Animal studies were performed in accordance with Animal Care and Use Committee guidelines of the University of Wisconsin-Madison.

**Biosafety and biosecurity.** All recombinant DNA protocols were approved by the University of Wisconsin-Madison's Institutional Biosafety Committee after risk assessments were conducted by the Office of Biological Safety, and by the University of Tokyo's Subcommittee on Living Modified Organisms, and, when required, by the competent minister of Japan. In addition, the University of Wisconsin-Madison Biosecurity Task Force regularly reviews the research program and ongoing activities of the laboratory. The task force has a diverse skill set and provides support in the areas of biosafety, facilities, compliance, security and health. Members of the Biosecurity Task Force are in frequent contact with the principal investigator and laboratory personnel to provide oversight and assure biosecurity. Experiments with viruses possessing the wild-type HA cleavage site were performed in enhanced BSL3 containment laboratories approved for such use by the CDC and the USDA. Ferret transmission studies were conducted by three scientists with both DVM and PhD degrees who each had more than a minimum of 6 years of experience with highly pathogenic influenza viruses and animal studies with highly pathogenic viruses. Our staff wear powered air-purifying respirators that filter the air, and disposable coveralls; they shower out on exit from the facility. The containment facilities at University of Wisconsin-Madison were designed to exceed standards outlined in Biosafety in Microbiological and Biomedical Laboratories (5th edition; <http://www.cdc.gov/biosafety/publications/bmbl5/BMbl.pdf>). Features of the BSL3-enhanced suites include entry/exit through a shower change room, effluent decontamination, negative air-pressure laboratories, double-door autoclaves, double HEPA-filtered exhaust air, and gas decontamination ports. The BSL3-Agriculture suite features include all those listed for BSL3-enhanced plus HEPA-filtered supply and double-HEPA-filtered exhaust air, double-gasketed watertight and airtight seals, airtight dampers on all ductwork, and the structure was pressure-decay tested during commissioning. The University of Wisconsin-Madison facility has a dedicated alarm system that monitors all building controls and sends alarms (~500 possible alerts). Redundancies and emergency resources are built-in to the facility including two



air handlers, two compressors, two filters each place filters are needed, two effluent sterilization tanks, two power feeds to the building, an emergency generator in case of a power failure and other physical containment measures in the facility that operate without power. Biosecurity monitoring of the facility is ongoing. All personnel undergo Select Agent security risk assessment by the United States Criminal Justice Information Services Division and complete rigorous biosafety, BSL3 and Select Agent training before participating in BSL3-level experiments. Refresher training is scheduled on a regular basis. The principal investigator participates in training sessions and emphasizes compliance to maintain safe operations and a responsible research environment. The laboratory occupational health plan is in compliance with the University of Wisconsin-Madison Occupational Health Program. Select agent virus inventory is checked monthly and submitted to the University of Wisconsin-Madison Research Compliance Specialist. Virus inventory is submitted 1–2 times per year to the file holder in the Select Agent branch of the CDC. The research program, procedures, occupational health plan, documentation, security and facilities are reviewed annually by the University of Wisconsin-Madison Responsible Official and at regular intervals by the CDC and the Animal and Plant Health Inspection Service (APHIS) as part of the University of Wisconsin-Madison Select Agent Program.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 18 August 2011; accepted 9 March 2012.**

**Published online 2 May 2012.**

- Glaser, L. *et al.* A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *J. Virol.* **79**, 11533–11536 (2005).
- Matrosovich, M. *et al.* Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J. Virol.* **74**, 8502–8512 (2000).
- Rogers, G. N. & Paulson, J. C. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* **127**, 361–373 (1983).
- Auewarakul, P. *et al.* An avian influenza H5N1 virus that binds to a human-type receptor. *J. Virol.* **81**, 9950–9955 (2007).
- Gambaryan, A. *et al.* Evolution of the receptor binding phenotype of influenza A (H5) viruses. *Virology* **344**, 432–438 (2006).
- Stevens, J. *et al.* Recent avian H5N1 viruses exhibit increased propensity for acquiring human receptor specificity. *J. Mol. Biol.* **381**, 1382–1394 (2008).
- Wang, W. *et al.* Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *J. Virol.* **84**, 6570–6577 (2010).
- Watanabe, Y. *et al.* Acquisition of human-type receptor binding specificity by new H5N1 influenza virus sublineages during their emergence in birds in Egypt. *PLoS Pathog.* **7**, e1002068 (2011).
- Yamada, S. *et al.* Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* **444**, 378–382 (2006).
- Jackson, S. *et al.* Reassortment between avian H5N1 and human H3N2 influenza viruses in ferrets: a public health risk assessment. *J. Virol.* **83**, 8131–8140 (2009).
- Maines, T. R. *et al.* Lack of transmission of H5N1 avian-human reassortant influenza viruses in a ferret model. *Proc. Natl Acad. Sci. USA* **103**, 12121–12126 (2006).
- Maines, T. R. *et al.* Effect of receptor binding domain mutations on receptor binding and transmissibility of avian influenza H5N1 viruses. *Virology* **413**, 139–147 (2011).
- Yen, H. L. *et al.* Inefficient transmission of H5N1 influenza viruses in a ferret contact model. *J. Virol.* **81**, 6890–6898 (2007).
- Chen, L. M. *et al.* *In vitro* evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology* **422**, 105–113 (2012).
- Stevens, J. *et al.* Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* **312**, 404–410 (2006).
- Neumann, G. *et al.* Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl Acad. Sci. USA* **96**, 9345–9350 (1999).
- Hatakeyama, S. *et al.* Enhanced expression of an  $\alpha$ 2,6-linked sialic acid on MDCK cells improves isolation of human influenza viruses and evaluation of their sensitivity to a neuraminidase inhibitor. *J. Clin. Microbiol.* **43**, 4139–4146 (2005).
- Shinya, K. *et al.* Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435–436 (2006).
- Varki, A. Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* **446**, 1023–1029 (2007).
- Rogers, G. N. *et al.* Host-mediated selection of influenza virus receptor variants. Sialic acid- $\alpha$ 2,6Gal-specific clones of A/duck/Ukraine/1/63 revert to sialic acid- $\alpha$ 2,3Gal-specific wild type *in ovo*. *J. Biol. Chem.* **260**, 7362–7367 (1985).
- Ha, Y., Stevens, D. J., Skehel, J. J. & Wiley, D. C. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proc. Natl Acad. Sci. USA* **98**, 11181–11186 (2001).
- Cline, T. D. *et al.* Increased pathogenicity of a reassortant 2009 pandemic H1N1 influenza virus containing an H5N1 hemagglutinin. *J. Virol.* **85**, 12262–12270 (2011).
- Octaviani, C. P., Ozawa, M., Yamada, S., Goto, H. & Kawaoka, Y. High level of genetic compatibility between swine-origin H1N1 and highly pathogenic avian H5N1 influenza viruses. *J. Virol.* **84**, 10918–10922 (2010).
- Li, H. *et al.* Isolation and characterization of H5N1 and H9N2 influenza viruses from pigs in China [in Chinese]. *Chin. J. Prev. Vet. Med.* **26**, 1–6 (2004).
- Nidom, C. A. *et al.* Influenza A (H5N1) viruses from pigs, Indonesia. *Emerg. Infect. Dis.* **16**, 1515–1523 (2010).
- Pasma, T. & Joseph, T. Pandemic (H1N1) 2009 infection in swine herds, Manitoba, Canada. *Emerg. Infect. Dis.* **16**, 706–708 (2010).
- Pereda, A. *et al.* Pandemic (H1N1) 2009 outbreak on pig farm, Argentina. *Emerg. Infect. Dis.* **16**, 304–307 (2010).
- Welsh, M. D. *et al.* Initial incursion of pandemic (H1N1) 2009 influenza A virus into European pigs. *Vet. Rec.* **166**, 642–645 (2010).
- Xu, Q., Wang, W., Cheng, X., Zengel, J. & Jin, H. Influenza H1N1 A/Solomon Island/3/06 virus receptor binding specificity correlates with virus pathogenicity, antigenicity, and immunogenicity in ferrets. *J. Virol.* **84**, 4936–4945 (2010).
- van Riel, D. *et al.* Human and avian influenza viruses target different cells in the lower respiratory tract of humans and other mammals. *Am. J. Pathol.* **171**, 1215–1223 (2007).
- Gao, Y. *et al.* Identification of amino acids in HA and PB2 critical for the transmission of H5N1 avian influenza viruses in a mammalian host. *PLoS Pathog.* **5**, e1000709 (2009).
- Itoh, Y. *et al.* *In vitro* and *in vivo* characterization of new swine-origin H1N1 influenza viruses. *Nature* **460**, 1021–1025 (2009).
- Skehel, J. J. & Wiley, D. C. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* **69**, 531–569 (2000).
- England, R. J., Homer, J. J., Knight, L. C. & Eli, S. R. Nasal pH measurement: a reliable and repeatable parameter. *Clin. Otolaryngol. Allied Sci.* **24**, 67–68 (1999).
- Carr, C. M., Chaudhry, C. & Kim, P. S. Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc. Natl Acad. Sci. USA* **94**, 14306–14313 (1997).
- Haywood, A. M. & Boyer, B. P. Time and temperature dependence of influenza virus membrane fusion at neutral pH. *J. Gen. Virol.* **67**, 2813–2817 (1986).
- Krenn, B. M. *et al.* Single HA2 mutation increases the infectivity and immunogenicity of a live attenuated H5N1 intranasal influenza vaccine candidate lacking NS1. *PLoS ONE* **6**, e18577 (2011).
- Treanor, J. J., Campbell, J. D., Zangwill, K. M., Rowe, T. & Wolff, M. Safety and immunogenicity of an inactivated subvirion influenza A (H5N1) vaccine. *N. Engl. J. Med.* **354**, 1343–1351 (2006).
- Bussey, K. A., Bousse, T. L., Desmet, E. A., Kim, B. & Takimoto, T. PB2 residue 271 plays a key role in enhanced polymerase activity of influenza A viruses in mammalian host cells. *J. Virol.* **84**, 4395–4406 (2010).
- Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* **293**, 1840–1842 (2001).
- Li, Z. *et al.* Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* **79**, 12058–12064 (2005).
- Mehle, A. & Doudna, J. A. Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc. Natl Acad. Sci. USA* **106**, 21312–21316 (2009).
- Yamada, S. *et al.* Biological and structural characterization of a host-adapting amino acid in influenza virus. *PLoS Pathog.* **6**, e1001034 (2010).
- Salzberg, S. L. *et al.* Genome analysis linking recent European and African influenza (H5N1) viruses. *Emerg. Infect. Dis.* **13**, 713–718 (2007).
- Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* **67**, 1761–1764 (1993).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors would like to acknowledge D. Holtzman for his contributions to the initial concept for this project and thoughtful scientific discussions. We thank M. McGregor, R. Moritz, L. Burley, K. Moore, A. Luka, J. Bettridge, N. Fujimoto and M. Ito for technical support, S. Watson for editing the manuscript, and the National Institute of Hygiene and Epidemiology, Hanoi, Vietnam for the A/Vietnam/1203/2004 (H5N1) virus, which was obtained from the CDC. This work was supported by the Bill & Melinda Gates Foundation (Grants 48339 and OPPGH5383), by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, by ERATO (Japan Science and Technology Agency), and by the National Institute of Allergy and Infectious Diseases Public Health Service Research grants. The following reagents were obtained from the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: polyclonal anti-monovalent influenza subvirion vaccine rgA/Vietnam/1203/2004 (H5N1), (antisera, Human), high titre pool, NR-4109 and low titre pool, NR-4110.

**Author Contributions** M.I., T.W., M.H., S.C.D., M.O., K.S., G.Z., A.H., H.K., S.W., C.L., S.Y., M.K., Y.S., E.A.M., G.N. and Y.K. designed the experiments; M.I., T.W., M.H., S.C.D., M.O., K.S., G.Z., A.H., H.K., S.W., C.L., S.Y. and M.K. performed the experiments; M.I., T.W., M.H., S.C.D., M.O., K.S., G.Z., A.H., H.K., S.W., C.L., E.K., S.Y., M.K., Y.S., E.A.M., G.N. and Y.K. analysed the data; M.I., T.W., M.H., S.C.D., K.S., E.A.M., G.N. and Y.K. wrote the manuscript; M.I., T.W. and M.H. contributed equally to this work.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to Y.K. ([kawaoka@svm.vetmed.wisc.edu](mailto:kawaoka@svm.vetmed.wisc.edu)).



## METHODS

**Cells.** Madin–Darby canine kidney (MDCK) cells and MDCK cells overexpressing Siaz2,6Gal (AX4 cells<sup>17</sup>) were maintained in Eagle's minimal essential medium (MEM) containing 5% newborn calf serum. Human embryonic kidney 293T cells were cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum (FBS). HeLa cells were maintained in MEM containing 10% FBS. All cells were maintained at 37 °C in 5% CO<sub>2</sub>.

**Plasmid construction and reverse genetics.** Plasmid constructs for viral RNA production (pPolI)—containing the genes of the A/Vietnam/1203/2004 (H5N1; VN1203), A/Puerto Rico/8/34 (H1N1; PR8), A/Kawasaki/173/2001 (H1N1; K173) and A/California/04/2009 (H1N1; CA04) viruses flanked by the human RNA polymerase I promoter and the mouse RNA polymerase I terminator—were constructed as described<sup>16</sup>. The multibasic amino acids at the haemagglutinin (HA) cleavage site (RERRRKKR↓G) of the reassortant viruses between VN1203 and PR8 were changed to RETR↓G by site-directed mutagenesis. All transfectant viruses were generated by using reverse genetics essentially as described previously<sup>16</sup>. Recombinant viruses were amplified in MDCK or AX4<sup>17</sup> cells and stored at –80 °C until use. The HA segment of all viruses was sequenced to ensure the absence of unwanted mutations. All experiments with the reassortant viruses between VN1203 and CA04 were performed in enhanced biosafety level 3 containment laboratories approved for such use by the CDC and the USDA.

To introduce random mutations into the globular head of the VN1203 HA protein, a 143-amino-acid region spanning residues 120–259 (H3 numbering) was selected. This region was subjected to PCR-based random mutagenesis by use of the GeneMorph II kit (Stratagene) following the manufacturer's instructions. The targeted mutation rate (1–2 amino acid replacements per molecule) was achieved through optimization of the template quantity, and was confirmed by sequence analysis of 48 individual clones. By using a PCR-based cloning strategy, we inserted the mutagenized region into its respective vector containing the VN1203 HA gene between the human RNA polymerase I promoter and mouse RNA polymerase I terminator sequences. The composition of the plasmid library was confirmed by sequencing. The plasmid library was then used to generate an influenza virus library, essentially as described<sup>16</sup>. The size of the virus library was  $7 \times 10^6$  p.f.u.

**Preparation of sialidase-treated TRBCs.** Turkey red blood cells (TRBCs) were washed three times with phosphate-buffered saline (PBS), and diluted to 20% (vol/vol) in PBS. TRBCs (1 ml) were incubated with 500 U of  $\alpha$ 2,3-sialidase from *Salmonella enterica* serovar Typhimurium LT2 (NEB) for 20–24 h at 37 °C, washed three times in PBS, and re-suspended in PBS or MEM containing 1% bovine serum albumin (BSA) (MEM/BSA).

**Haemagglutination assay.** Viruses (50  $\mu$ l) were serially diluted with 50  $\mu$ l of PBS in a microtitre plate. An equal volume (that is, 50  $\mu$ l) of a 0.5% (vol/vol) TRBC suspension was added to each well. The plates were kept at room temperature and haemagglutination was assessed after a 1-h incubation.

**Virus library screening.** To select VN1203 HA variants that had acquired the ability to recognize human-type receptors, three parallel experiments were carried out, each with  $0.7 \times 10^6$  viruses. The virus library was first incubated with 0.1 ml of 10% (vol/vol)  $\alpha$ 2,3-sialidase-treated TRBCs for 10 min at 4 °C. After this incubation, the TRBCs and bound viruses were pelleted at 1,000 r.p.m. for 1 min, and the pellets then washed ten times in MEM/BSA containing 313 mM NaCl. Bound viruses were eluted by incubation at 37 °C for 30 min and then diluted to approximately 0.5 virus per well (determined by virus titration in a pilot study). Individual viruses were then amplified in AX4 cells, which overexpress Siaz2,6Gal<sup>17</sup>. Individual viruses were re-screened by using haemagglutination assays with  $\alpha$ 2,3-sialidase-treated TRBCs.

**Solid-phase binding assay.** Viruses were grown in MDCK cells, clarified by low-speed centrifugation, laid over a cushion of 30% sucrose in PBS, and ultracentrifuged at 25,000 r.p.m. for 2 h at 4 °C. Virus stocks were aliquoted and stored at –80 °C. Virus concentrations were determined by using haemagglutination assays with 0.5% (vol/vol) TRBCs. The direct receptor-binding capacity of viruses was examined by use of a solid-phase binding assay as previously described<sup>9</sup>. Microtitre plates (Nunc) were incubated with the sodium salts of sialylglycopolymers (poly-L-glutamic acid backbones containing N-acetylneuraminic acid linked to galactose through either an  $\alpha$ 2,3 (Neu5Ac2,3Gal $\beta$ 1,4GlcNAc $\beta$ 1-pAP) or an  $\alpha$ 2,6 (Neu5Ac2,6Gal $\beta$ 1,4GlcNAc $\beta$ 1-pAP) bond) in PBS at 4 °C overnight. After the glycopolymer solution was removed, the plates were blocked with 0.15 ml of PBS containing 4% BSA at room temperature for 1 h. After four successive washes with ice-cold PBS, the plates were incubated in a solution containing influenza virus (8–32 HA units in PBS) at 4 °C overnight. After washing as described above, the plates were incubated for 2 h at 4 °C with rabbit polyclonal antiserum to either K173 or VN1203 virus. The plates were then washed again as before and incubated with horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG antiserum for 2 h at 4 °C. After washing, the plates were incubated with O-phenylenediamine

(Sigma) in PBS containing 0.01% H<sub>2</sub>O<sub>2</sub> for 10 min at room temperature, and the reaction was stopped with 0.05 ml of 1 M HCl. The optical density at 490 nm was determined in a plate reader (Infinite M1000; Tecan).

**Virus binding to human airway tissues.** Paraffin-embedded normal human trachea (US Biological) and lung (BioChain) tissue sections were deparaffinized and rehydrated. Sections were then blocked by using 4% BSA in PBS and covered with virus suspensions (64 HA units in PBS) at 4 °C overnight. After being washed four times in ice-cold PBS, the sections were incubated with primary antibodies for 3 h at 4 °C. The primary antibodies used were as follows: a pool of mouse anti-VN1203 HA monoclonal antibodies (15A3, 3G2, 7A11, 8A3, 14C5 and 18E1; Rockland); rabbit anti-K173 polyclonal antibody; rabbit anti-surfactant protein A polyclonal antibody (Millipore); and mouse anti-surfactant protein A monoclonal antibody (Abcam). Antibody binding was detected by using an IgG secondary antibody conjugated with Alexa Fluor 488 or Alexa Fluor 633 (Molecular Probes). Sections were also counterstained with Hoechst 33342, trihydrochloride, trihydrate (Molecular Probes). The samples were examined by using confocal laser scanning microscopy (model LSM 510; Carl Zeiss).

To confirm sialic-acid-specific virus binding, tissue sections were treated, before incubation with viruses, with *Arthrobacter ureafaciens* sialidase (Sigma) for 3 h at 37 °C. Viruses bound to tissue were detected as described above.

**Experimental infection of ferrets.** Animal studies were performed in accordance with the Animal Care and Use Committee guidelines of the University of Wisconsin–Madison. We used 6–10-month-old female ferrets (Triple F Farms) that were serologically negative by haemagglutination inhibition (HI) assay for currently circulating human influenza viruses. Six ferrets per group were anaesthetized intramuscularly with ketamine and xylazine (5–30 mg and 0.2–6 mg kg<sup>–1</sup> of body weight, respectively) and inoculated intranasally with  $10^6$  p.f.u. (500  $\mu$ l) of viruses. On days 3 and 6 after infection, three ferrets per group were killed for virological and pathological examinations. The virus titres in various organs were determined by use of plaque assays in MDCK cells.

Excised tissue samples of nasal turbinates, trachea, lungs, brain, liver, spleen, kidney and colon from euthanized ferrets were preserved in 10% phosphate-buffered formalin. Tissues were then trimmed and processed for paraffin embedding and cut into 5- $\mu$ m-thick sections. One section from each tissue sample was stained by using a standard haematoxylin-and-eosin procedure, whereas another one was processed for immunohistological staining with a mixture of two anti-influenza virus rabbit antibodies (1:2,000; R309 and anti-VN1203; both prepared in our laboratory) that react with CA04 and VN1203, respectively. Specific antigen–antibody reactions were visualized by using an indirect two-step dextran-polymer technique (Dako EnVision system; Dako) and 3,3'-diaminobenzidine tetrahydrochloride staining (Dako).

**Ferret transmission study.** For transmission studies in ferrets, animals were housed in adjacent transmission cages that prevented direct and indirect contact between animals but allowed spread of influenza virus through the air (Showa Science; Supplementary Fig. 7). Three, five, or six ferrets were inoculated intranasally with  $10^6$  p.f.u. (500  $\mu$ l) of virus (inoculated ferrets). Twenty-four hours after infection, three, five, or six naive ferrets were each placed in a cage adjacent to an inoculated ferret (contact ferrets). The ferrets were monitored for changes in body weight and the presence of clinical signs. To assess viral replication in nasal turbinates, we determined viral titres in nasal washes collected from virus-inoculated and contact ferrets on day 1 after inoculation or co-housing, respectively, and then every other day.

**Serological tests.** Serum samples were collected between days 14 and 20 after infection, treated with receptor-destroying enzyme, heat-inactivated at 56 °C for 30 min, and tested by use of an HI assay with 0.5% TRBCs ([http://www.wpro.who.int/entity/emerging\\_diseases/documents/docs/manualanimalaidiagnosisandsurveillance.pdf](http://www.wpro.who.int/entity/emerging_diseases/documents/docs/manualanimalaidiagnosisandsurveillance.pdf)). Viruses bearing homologous HA were used as antigens for the HI tests.

**Polykaryon formation representing membrane fusion activity.** Monolayers of HeLa cells grown in 12-well plates were transfected with the protein expression vector pCAGGS<sup>46</sup> encoding wild-type or mutant HA. At 24 h after transfection, cells transiently expressing HA protein were treated with trypsin (1  $\mu$ g ml<sup>–1</sup>) in MEM containing 0.3% BSA for 30 min at 37 °C to cleave the HA into its HA1 and HA2 subunits. Polykaryon formation was induced by exposing the cells to low-pH buffer (145 mM NaCl, 20 mM sodium citrate (pH 6.0–5.4)) for 2 min at 37 °C. After this exposure, the low-pH buffer was replaced with MEM containing 10% FBS and the cells were incubated for 3 h at 37 °C. The cells were then fixed with methanol and stained with Giemsa's solution and photographed with a digital camera mounted on an inverted microscope (Nikon, Eclipse Ti). For quantitative analyses, cell nuclei were counted in five randomly chosen fields of cell culture. Polykaryon formation activity was calculated from the number of nuclei in polykaryons divided by the total number of nuclei in the same field.

**Thermostability.** Viruses (128 HA units in PBS) were incubated for the times indicated at 50 °C. Subsequently, infectivity and haemagglutination activity were determined by use of plaque assays in MDCK cells and haemagglutination assays using 0.5% TRBCs, respectively.

**Neuraminidase (NA) inhibition assay.** To assess the sensitivity of viruses to the NA inhibitor oseltamivir, NA inhibition assays were performed as described previously<sup>32</sup>.

**Statistical analysis.** All statistical analyses were performed using JMP 9.0.0 (SAS Institute Inc.). The statistical significance of differences between rgCA04 and H5 avian/human reassortant viruses was determined by using a Dunnett's test. Comparisons of polykaryon formation between wild-type and mutant HAs were done using Tukey's test. *P* values of <0.05 were considered significant.

46. Niwa, H., Yamamura, K. & Miyazaki, J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193–199 (1991).

# Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly

Yong-Chun Yu<sup>1\*</sup>, Shuijin He<sup>2\*</sup>, She Chen<sup>2</sup>, Yinghui Fu<sup>1</sup>, Keith N. Brown<sup>2,3</sup>, Xing-Hua Yao<sup>1</sup>, Jian Ma<sup>1</sup>, Kate P. Gao<sup>2,3</sup>, Gina E. Sosinsky<sup>4</sup>, Kun Huang<sup>5</sup> & Song-Hai Shi<sup>2,3</sup>

Radial glial cells are the primary neural progenitor cells in the developing neocortex<sup>1</sup>. Consecutive asymmetric divisions of individual radial glial progenitor cells produce a number of sister excitatory neurons that migrate along the elongated radial glial fibre, resulting in the formation of ontogenetic columns<sup>2–4</sup>. Moreover, sister excitatory neurons in ontogenetic columns preferentially develop specific chemical synapses with each other rather than with nearby non-siblings<sup>5</sup>. Although these findings provide crucial insight into the emergence of functional columns in the neocortex, little is known about the basis of this lineage-dependent assembly of excitatory neuron microcircuits at single-cell resolution. Here we show that transient electrical coupling between radially aligned sister excitatory neurons regulates the subsequent formation of specific chemical synapses in the neocortex. Multiple-electrode whole-cell recordings showed that sister excitatory neurons preferentially form strong electrical coupling with each other rather than with adjacent non-sister excitatory neurons during early postnatal stages. This preferential coupling allows selective electrical communication between sister excitatory neurons, promoting their action potential generation and synchronous firing. Interestingly, although this electrical communication largely disappears before the appearance of chemical synapses, blockade of the electrical communication impairs the subsequent formation of specific chemical synapses between sister excitatory neurons in ontogenetic columns. These results suggest a strong link between lineage-dependent transient electrical coupling and the assembly of precise excitatory neuron microcircuits in the neocortex.

Sister excitatory neurons in individual ontogenetic columns preferentially develop specific chemical synapses with each other rather than with adjacent non-sister excitatory neurons<sup>5</sup>. Given the almost complete overlap of the dendritic fields of neighbouring excitatory neurons, it is unclear how this lineage-dependent assembly of precise columnar microcircuits is controlled at the individual cell level. Some studies have suggested that gap-junction-mediated neuronal communication is involved in the formation of local connectivity in the developing neocortex<sup>6–8</sup>, even though direct evidence of electrically coupled neocortical neurons at early developmental stages is lacking. In this study, we set out to investigate whether gap-junction-mediated electrical coupling exists between sister excitatory neurons in ontogenetic columns and, if so, whether this coupling regulates the preferential formation of chemical synapses between sister excitatory neurons.

Gap junctions are composed of two membrane-docked hexameric hemi-channels that consist of connexin proteins from two adjacent cells. There are ~20 genes encoding connexins in rodents, and the corresponding protein symbols are denoted as CX plus the calculated molecular mass of the protein<sup>9</sup>. Of these proteins, CX26 and CX43 have been shown to be abundantly expressed in the developing neocortex at embryonic and neonatal stages<sup>10,11</sup>. Consistent with this,

we found that developing neurons in the neonatal neocortex expressed CX26 (Supplementary Fig. 1a). Moreover, CX26-positive puncta were present at the dendrodendritic and dendrosomatic contacts of radially aligned sister excitatory neurons that were labelled by *in utero* intraventricular injection of low-titre enhanced green fluorescence protein (eGFP)-expressing retrovirus at embryonic day 12 to 13 (E12–13) (Fig. 1a–e and Supplementary Movie 1), indicating the existence of gap junctions between sister excitatory neurons in ontogenetic columns.

Gap junctions mediate intercellular adhesion and the exchange of small molecules (typically less than 1 kDa), including low-molecular-mass dyes and ions that can be detected experimentally<sup>12,13</sup>. Previous dye injection experiments have suggested the presence of gap junctions between progenitor cells in the embryonic neocortex<sup>2,14</sup> and between neurons in the neonatal neocortex<sup>6,15</sup>. Although these studies have provided important insight, the accuracy of dye coupling in revealing the presence of gap junctions has been debated<sup>12,13</sup>. To circumvent this issue and to quantitatively examine gap junction channel activity, we performed whole-cell patch-clamp recording experiments to study gap junctions between sister excitatory neurons in ontogenetic columns.

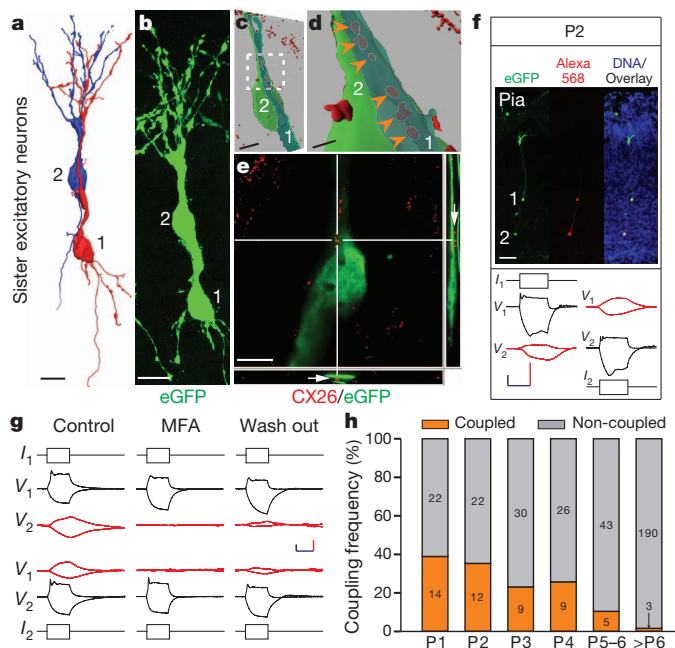
We prepared acute neocortical slices from postnatal mice (postnatal day 1 to 28, P1–28) that had received *in utero* intraventricular injection of eGFP-expressing retrovirus at E12–13. Guided by infrared differential interference contrast (DIC) and epifluorescence illumination, we simultaneously recorded from two radially aligned eGFP-expressing sister excitatory neurons (Fig. 1f and Supplementary Fig. 1b, top (green)). We identified excitatory neurons on the basis of their morphological characteristics, including a pyramid-shaped cell body and a long apical dendrite. After the recordings were established, we tested the electrical coupling between the neurons under current-clamp conditions (Fig. 1f and Supplementary Fig. 1b, bottom) or voltage-clamp conditions (Supplementary Fig. 2a). Under current-clamp conditions, hyperpolarization ( $V_1$ ) of one of the neurons (neuron 1, the driver) by current injection ( $I_1$ ) produced a simultaneous hyperpolarization ( $V_2$ ) of the non-injected neuron (neuron 2, the receiver) (Fig. 1f and Supplementary Fig. 1b, bottom). Similarly, depolarization of the driver neuron (neuron 1) produced a simultaneous depolarization of the receiver neuron (neuron 2). As expected for electrotonic propagation, voltage deflections recorded in the non-injected receiver neuron (Fig. 1f, red, and Supplementary Fig. 1b, bottom) had a smaller amplitude and a slower time course than those in the injected driver neuron (Fig. 1f, black, and Supplementary Fig. 1b, bottom). In all cases, electrical transmission between the two neurons was found to be reciprocal (Fig. 1f and Supplementary Fig. 1b, bottom). Similar bidirectional electrotonic propagation was recorded under voltage-clamp conditions (Supplementary Fig. 2a).

To confirm that the electrical coupling between sister excitatory neurons is mediated by gap junctions, we exposed electrically coupled sister excitatory neuron pairs to the gap junction blocker meclofenamic

<sup>1</sup>Institute of Neurobiology, Institutes of Brain Science and State Key Laboratory of Medical Neurobiology, Fudan University, 138 Yixueyuan Road, Shanghai 200032, China. <sup>2</sup>Developmental Biology Program, Memorial Sloan-Kettering Cancer Centre, 1275 York Avenue, New York, New York 10065, USA. <sup>3</sup>Neuroscience Graduate Program, Weill Cornell Medical College, 1230 York Avenue, New York, New York 10065, USA. <sup>4</sup>National Centre for Microscopy and Imaging Research and Department of Neurosciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0608, USA. <sup>5</sup>Department of Biomedical Informatics, Comprehensive Cancer Center Biomedical Informatics Shared Resource, The Ohio State University, 333 West 10th Avenue, Columbus, Ohio 43210, USA.

\*These authors contributed equally to this work.





**Figure 1 | Gap-junction-mediated electrical coupling between sister excitatory neurons in neonatal neocortical ontogenetic columns.**

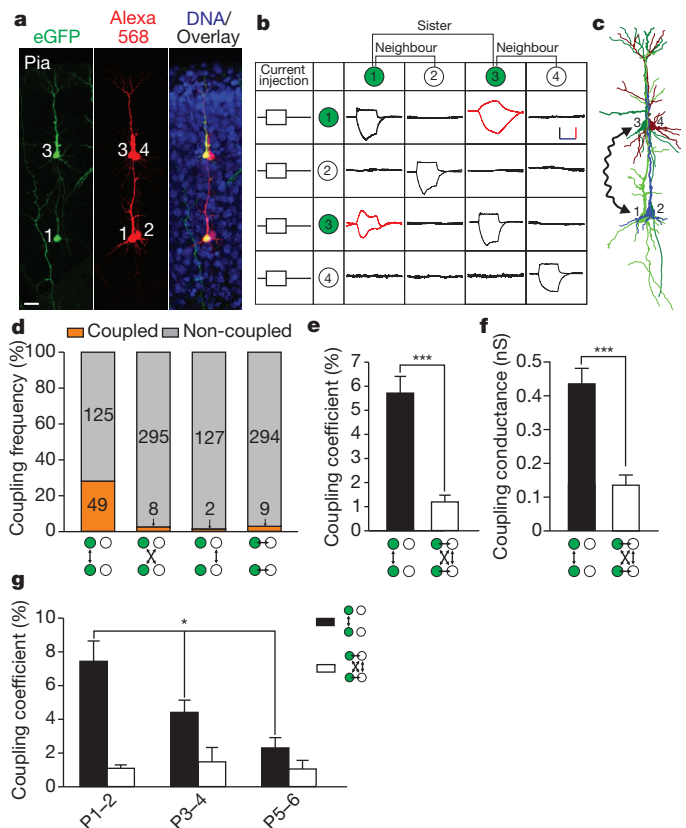
**a–e**, Images and three-dimensional reconstruction of two radially aligned eGFP-expressing sister excitatory neurons (1 and 2) (green) immunostained with antibody specific for CX26 (red). (A morphological reconstruction is shown in **a**, in which neuron 1 is shown in red and neuron 2 in blue.) Note the presence of CX26 puncta at the dendrosomatic and dendrodendritic contacts of sister excitatory neurons in the three-dimensional reconstruction (**c**) and indicated by arrowheads in **d**, which shows a magnification of the boxed region in **c**) and in the Z-axis cross section (**e**, arrows). Scale bars, 10 µm (**a**, **b**); 5 µm (**c**); 1 µm (**d**); and 5 µm (**e**). **f**, Dual whole-cell recordings of sister excitatory neurons in ontogenetic columns at the neonatal stage P2. Top, images of sister excitatory neurons expressing eGFP (green) in an ontogenetic column; the neurons have been filled with Alexa Fluor 568 hydrazide (red) through the recording pipette and stained with the DNA-binding dye 4',6-diamidino-2-phenylindole (DAPI; blue). Scale bar, 50 µm. Bottom, simultaneous depolarization or hyperpolarization of two sister excitatory neurons ( $V_1$  and  $V_2$ ) when a positive or negative current (+100 pA or -100 pA) was injected into only one of the neurons ( $I_1$  or  $I_2$ ) under current-clamp conditions, indicating electrical coupling. Subsequent figures use a similar panel layout. Scale bars, 50 mV (black), 10 mV (red) and 200 ms (blue). Pia, pial surface. **g**, Blockade of electrical coupling by the gap junction blocker MFA (100 µM). Scale bars, 25 mV (black), 5 mV (red) and 200 ms (blue). **h**, Summary of the frequency of electrical coupling between sister excitatory neurons in ontogenetic columns at different postnatal stages.

acid (MFA, 100 µM), the glutamate receptor antagonists 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX, 10 µM) and 3-(2-carboxypiperazin-4-yl)propyl-1-phosphonic acid (CPP, 20 µM), or the GABA<sub>A</sub> (γ-aminobutyric acid A) receptor antagonist bicuculline methiodide (BMI, 10 µM). Neither BMI nor CNQX plus CPP had any effect ( $n = 3$ ; Supplementary Fig. 2b); however, MFA completely eliminated electrical transmission in all cases ( $n = 9$ ; Fig. 1g). Moreover, electrical transmission recovered partially after MFA was washed out (Fig. 1g).

The frequency of observed gap-junction-mediated electrical coupling between sister excitatory neurons in ontogenetic columns at P1 and P2 was 38.9% (occurring in 14 of the 36 pairs tested) and 35.3% (12 of 34 pairs tested), respectively (Fig. 1h). As development proceeded, this coupling frequency progressively decreased to ~20–25% at P3–P4 (9 of 39 pairs tested at P3 and 9 of 35 pairs tested at P4) and to ~10% at P5–P6 (5 of 48 pairs tested) (Fig. 1h). After P6, electrical coupling between sister excitatory neurons was rare; only 3 of 193 pairs were coupled (Fig. 1h); this finding is consistent with previous observations of very sparse electrical coupling among more mature excitatory neurons in the neocortex<sup>16–18</sup>. No chemical synapses were detected between

sister excitatory neurons at early postnatal stages (P1–P6), as previously reported<sup>5</sup>.

To test whether sister excitatory neurons preferentially form gap junctions with each other, we performed quadruple whole-cell recordings on two eGFP-expressing sister excitatory neurons in individual ontogenetic columns and on two non-eGFP-expressing excitatory neurons adjacent to the eGFP-expressing neurons on the same side in the developing neocortex (Fig. 2a). We carried out this experiment at P1–P6, when electrical coupling is prominent (Fig. 1h). Once all four recordings were established, hyperpolarizing and depolarizing currents were injected sequentially into one of the four neurons, and the voltage changes were monitored in all four neurons to probe gap-junction-mediated electrical coupling. When eGFP-expressing neuron 1 was hyperpolarized or depolarized, only its sister neuron



**Figure 2 | Preferential formation of strong electrical coupling between sister excitatory neurons in neonatal neocortical ontogenetic columns.**

**a**, Image of a quadruple whole-cell recording of two eGFP-expressing sister excitatory neurons (neurons 1 and 3, green) in an ontogenetic column and two non-eGFP-expressing excitatory neurons (neurons 2 and 4) adjacent to the sisters on the same side, filled with Alexa Fluor 568 hydrazide (red) and stained with DAPI (blue). Scale bar, 20 µm. **b**, Sample traces of voltage changes in the four neurons in response to sequential current injection into one of the four neurons. Green circles indicate eGFP-expressing sister excitatory neurons, and white circles indicate non-eGFP-expressing neighbouring excitatory neurons. The average traces are shown in each table cell. Scale bars, 50 mV (black), 5 mV (red) and 200 ms (blue). **c**, A morphological reconstruction of the four neurons in the quadruple recording. The wavy arrow indicates reciprocal electrical coupling between sister excitatory neurons 1 and 3. **d**, Summary of the frequency of electrical coupling observed between sister excitatory neurons in ontogenetic columns and their adjacent non-sister excitatory neurons at P1–P6. **e**, **f**, Summary of the coupling coefficient (**e**) (sisters,  $n = 74$ ; non-sisters,  $n = 24$ ; \*\*\* $P < 1 \times 10^{-8}$ ; data are mean  $\pm$  s.e.m.) and conductance (**f**) (sisters,  $n = 85$ ; non-sisters,  $n = 24$ ; \*\*\* $P < 5 \times 10^{-7}$ ; data are mean  $\pm$  s.e.m.) at P1–P6. **g**, Progressive decrease in the coupling coefficient between sister excitatory neurons in ontogenetic columns as development proceeds (sisters, P1–P2,  $n = 36$ ; P3–P4,  $n = 32$ ; P5–P6,  $n = 6$ ; \* $P < 0.01$ ; non-sisters, P1–P2,  $n = 11$ ; P3–P4,  $n = 7$ ; P5–P6,  $n = 6$ ;  $P = 0.7$ ; data are mean  $\pm$  s.e.m.).

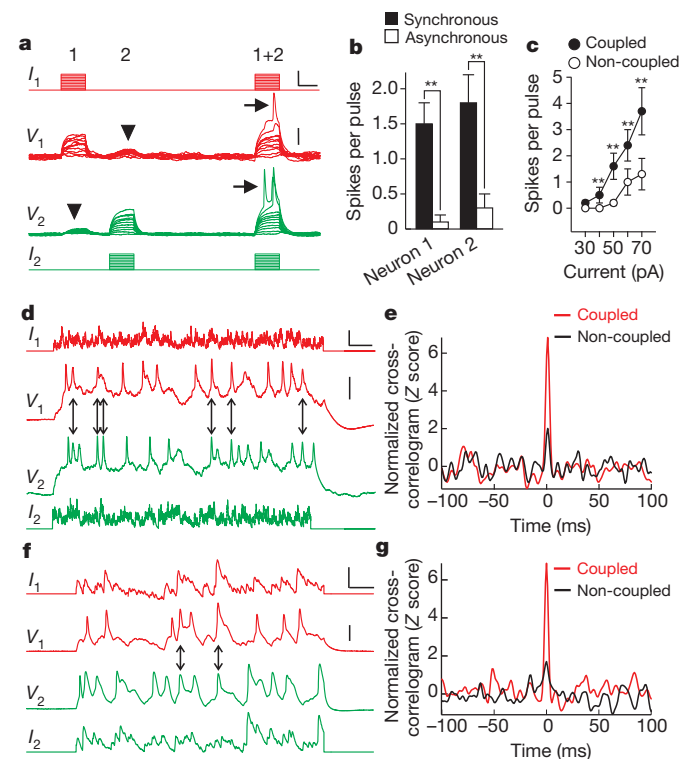
(neuron 3) showed simultaneous hyperpolarization or depolarization (Fig. 2b), despite the almost complete overlap of the neurite arbours of the adjacent neurons, neurons 3 and 4 (Fig. 2c). Similarly, when eGFP-expressing neuron 3 was hyperpolarized or depolarized, only its sister neuron (neuron 1) showed simultaneous hyperpolarization or depolarization (Fig. 2b). The hyperpolarization or depolarization of neuron 2 or 4 failed to trigger simultaneous voltage deflection in any of the other three neurons (Fig. 2b).

We analysed a total of 174 pairs of radially aligned eGFP-expressing sister excitatory neurons and their neighbouring non-sibling excitatory neurons at P1–P6 (Fig. 2d). Of the sister excitatory neuron pairs in an ontogenetic column, 28.2% (49 of 174 pairs) were electrically coupled. By contrast, only 2.6% (8 of 303 pairs) of radially situated non-sister excitatory neuron pairs (one eGFP-expressing and one non-eGFP-expressing) were electrically coupled (Fig. 2d). In addition, only 1.6% (2 of 129 pairs) of similarly radially situated non-eGFP expressing excitatory neuron pairs were coupled, and only 3.0% (9 of 303) of nearby eGFP-expressing and non-eGFP-expressing excitatory neuron pairs were coupled (Fig. 2d). These findings are consistent with a recent study showing almost no electrical coupling between randomly selected excitatory neurons in the developing neocortex<sup>19</sup>. Our results clearly demonstrate that sister excitatory neurons in ontogenetic columns have a strong preference for gap-junction-mediated electrical coupling with each other rather than with adjacent non-sister excitatory neurons and that the frequency of electrical coupling between non-sister excitatory neurons in the developing neocortex is low.

We also compared the coupling coefficient between coupled sister excitatory neuron pairs and rarely coupled non-sister excitatory neuron pairs. The coupling coefficient, estimated as the ratio of the amplitude of the low-frequency voltage change in the receiver neuron to that in the driver neuron, reflects the strength of the electrical coupling. We found that the coupling coefficient between coupled sister excitatory neurons was substantially higher than that between coupled non-sister excitatory neurons (sisters,  $5.7 \pm 0.7\%$ ,  $n = 74$ ; non-sisters,  $1.2 \pm 0.3\%$ ,  $n = 24$ ;  $P < 1 \times 10^{-8}$ ; Fig. 2e). We also estimated the coupling conductance under voltage-clamp conditions and found that the coupling conductance between sister excitatory neurons was significantly larger than that between non-sister excitatory neurons (sisters,  $0.43 \pm 0.05$  nS, range 0.06–2.16 nS,  $n = 85$ ; non-sisters,  $0.13 \pm 0.03$  nS, range 0.06–0.81 nS,  $n = 24$ ;  $P < 5 \times 10^{-7}$ ; Fig. 2f). These results suggest that the electrical coupling between sister excitatory neurons is much stronger than that between non-sisters. Furthermore, we observed a progressive decrease in the coupling coefficient of coupled sister excitatory neuron pairs as development proceeded (P1–P2,  $7.4 \pm 1.2\%$ ,  $n = 36$ ; P3–P4,  $4.4 \pm 0.7\%$ ,  $n = 32$ ; P5–P6,  $2.3 \pm 0.6\%$ ,  $n = 6$ ;  $P < 0.01$ ), whereas the coupling coefficient of rarely coupled non-sister pairs did not change significantly during this period (P1–P2,  $1.1 \pm 0.2\%$ ,  $n = 11$ ; P3–P4,  $1.5 \pm 0.8\%$ ,  $n = 7$ ; P5–P6,  $1.1 \pm 0.5\%$ ,  $n = 6$ ;  $P = 0.7$ ; Fig. 2g).

Having found that sister excitatory neurons in ontogenetic columns preferentially form strong gap-junction-mediated electrical coupling with each other, we then examined the properties of this electrical transmission and tested whether this selective electrical coupling modulates the neuronal activity of sister excitatory neurons at neonatal stages. We injected a series of sinusoidal current waveforms of the same amplitude but different frequencies into one excitatory neuron and measured the response in its coupled sister excitatory neuron (Supplementary Fig. 3). We found that responses to low-frequency sine waves showed higher coupling coefficients and smaller phase lags than those triggered by high-frequency sine waves (Supplementary Fig. 3). These results suggest that the efficacy of signal transmission through electrical coupling between sister excitatory neurons is frequency dependent, similarly to previous observations of electrical coupling between inhibitory interneurons in the more mature neocortex<sup>16,17</sup>.

To determine whether this selective electrical coupling modulates neuronal activity, we injected subthreshold depolarizing current pulses at the same time or at different times into two eGFP-labelled, electrically coupled sister excitatory neurons in ontogenetic columns (Fig. 3a). In most cases, the asynchronous pulses did not generate an action potential in either neuron (Fig. 3a, arrowheads, and Fig. 3b, open bars; neuron 1,  $0.1 \pm 0.1$  spikes per pulse; neuron 2,  $0.3 \pm 0.2$  spikes per pulse). However, when the same current pulses were synchronously injected into both neurons, the two sister neurons reached action potential threshold and generated spikes (Fig. 3a, arrows, and Fig. 3b, filled bars; neuron 1,  $1.5 \pm 0.3$  spikes per pulse; neuron 2,  $1.8 \pm 0.4$  spikes per pulse). Similar observations were made in seven pairs of electrically coupled sister excitatory neurons but not in non-coupled sister neuron pairs (Fig. 3c;  $P < 0.001$ ). These results show that selective electrical coupling can strongly facilitate the generation of action potentials in coupled sister excitatory neurons in ontogenetic columns.



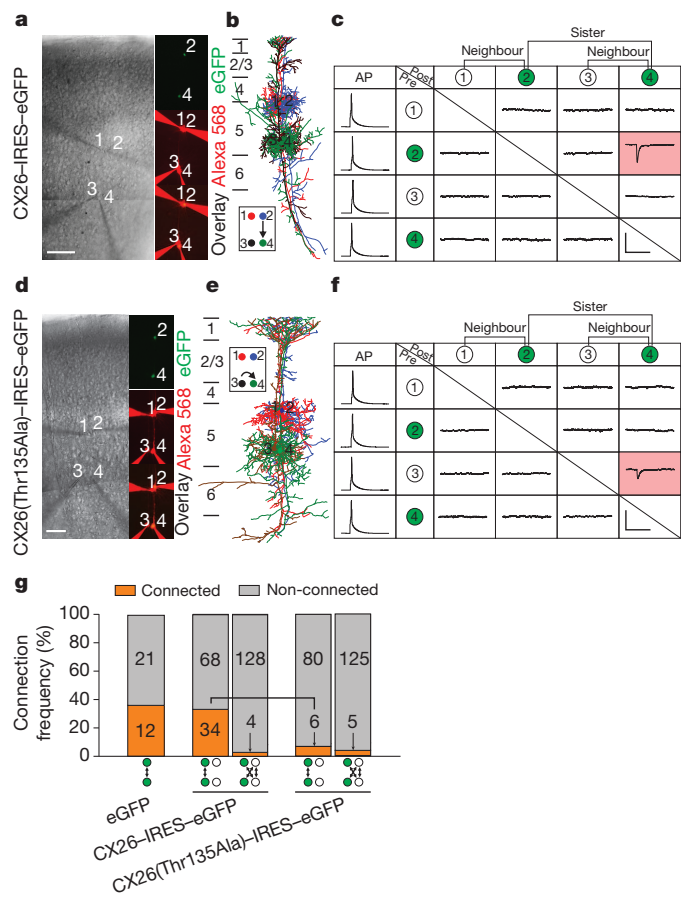
**Figure 3 | Electrical coupling promotes action potential generation and synchronous firing of sister excitatory neurons in neonatal neocortical ontogenetic columns.** **a–c**, Electrical coupling facilitates action potential generation in sister excitatory neurons. **a**, Sample traces of synchronous (1 + 2) or asynchronous (1 or 2) injection of subthreshold current pulses into electrically coupled sister excitatory neurons in an ontogenetic column. Arrowheads indicate voltage deflection due to the electrical transmission. Note that synchronous injection, but not asynchronous injection, results in action potential generation (arrows). Scale bars, 50 pA, 20 mV and 100 ms. **b**, Summary of the firing rate of the two sister excitatory neurons in **a** responding to a 50-pA current injection ( $**P < 0.001$ ; data are mean  $\pm$  s.e.m.). **c**, Summary of the firing rate in electrically coupled or non-coupled sister excitatory neurons responding to simultaneous current injections ( $n = 7$ ;  $**P < 0.001$ ; data are mean  $\pm$  s.e.m.). **d–g**, Electrical coupling promoting synchronous firing of sister excitatory neurons in ontogenetic columns in response to uncorrelated simulated neuronal activity (**d**, **e**) or uncorrelated native neuronal activity (**f**, **g**). **d**, **f**, Sample traces of voltage changes in electrically coupled sister excitatory neurons. Arrows indicate the spikes that occur in both neurons within a 1-ms window. Scale bars, 200 pA, 30 mV and 100 ms (**d**); and 100 pA, 25 mV and 100 ms (**f**). **e**, **g**, Normalized cross-correlogram analysis. The bin size is 1 ms. Note that the firing frequency is significantly increased near 0 ms for coupled sister excitatory neuron pairs (red) but not for non-coupled sister excitatory neuron pairs (black), indicating synchronous firing.

We next determined whether this selective electrical coupling could facilitate synchronous spiking of radially aligned sister excitatory neurons in response to natural stimuli. First, we simulated natural activity by injecting two electrically coupled or non-coupled sister excitatory neurons with uncorrelated random current signals obtained by convolving the Poisson trains of recorded spontaneous excitatory postsynaptic current waveforms<sup>17,20</sup> ( $I_1$  and  $I_2$ ; Fig. 3d). Second, we directly recorded the spontaneous subthreshold activity of neurons in the neonatal neocortex and then injected two electrically coupled or non-coupled sister excitatory neurons with signals corresponding to the recorded uncorrelated native neuronal activity ( $I_1$  and  $I_2$ ; Fig. 3f). We then analysed the cross-correlogram of firing induced in these sister excitatory neurons and found that there was a significant increase in firing centred at a time of 0 ms in coupled pairs compared with non-coupled ones (Fig. 3e, coupled,  $n = 4$ ; non-coupled,  $n = 5$ ;  $P < 0.005$ ; and Fig. 3g, coupled,  $n = 10$ ; non-coupled,  $n = 9$ ;  $P < 0.05$ ). These results suggest that gap-junction-mediated electrical transmission between sister excitatory neurons in ontogenetic columns can effectively promote precise (within 1 ms) synchronous firing in response to uncorrelated neuronal activity.

It has long been postulated that correlated neuronal activity facilitates chemical synapse formation and neuronal circuit assembly<sup>21,22</sup>. Given that sister excitatory neurons in ontogenetic columns preferentially form strong gap-junction-mediated electrical coupling, which promotes action potential generation and synchronous firing, this process could provide a mechanism by which chemical synapses form preferentially between sister excitatory neurons in ontogenetic columns. We therefore tested whether the selective electrical coupling between sister excitatory neurons is required for preferential formation of chemical synapses within individual ontogenetic columns.

Our immunohistochemistry experiments suggest that CX26 is a major connexin isoform that mediates electrical coupling between sister excitatory neurons in the developing neocortex (Fig. 1a–e and Supplementary Fig. 1a). Previous studies have shown that mutation of a conserved threonine residue (Thr 135) in the third transmembrane helix of CX26 to an alanine residue (CX26(Thr135Ala)) creates dominant-negative closed gap junction channels without affecting the synthesis, assembly or trafficking of the channels<sup>23</sup>. Mutant CX26 also exerts a *trans*-dominant-negative effect on other connexins<sup>24</sup>, providing a useful tool for eliminating broad gap junction channel functionality<sup>23</sup>. Notably, this point mutation should not interfere with gap-junction-mediated adhesion that is required for the proper radial migration of newly generated excitatory neurons in the developing neocortex<sup>10</sup>. We therefore engineered retroviruses expressing either wild-type CX26 or the dominant-negative closed channel mutant CX26(Thr135Ala), together with eGFP, using an internal ribosomal entry site (IRES) sequence, and performed *in utero* intraventricular injection of the retroviruses at E12–E13.

We found that expression of wild-type CX26 had no discernible effect, but expression of CX26(Thr135Ala) largely eliminated the electrical coupling between sister excitatory neurons at P1–P5 (CX26, coupling frequency 26.4%, 14 of 53 pairs; CX26(Thr135Ala), 9.8%, 4 of 41 pairs; Supplementary Fig. 4). More importantly, when we examined the chemical synapses formed between sister excitatory neurons at P10–P21, we found that only ~7.0% (6 of 86) of the sister pairs expressing the dominant-negative closed channel mutant CX26(Thr135Ala) were connected by chemical synapses (Fig. 4d–g), compared to 33–36% of the sister pairs expressing either eGFP alone (12 of 33 pairs tested) or wild-type CX26 (34 of 102 pairs tested) (Fig. 4a–c, g). The expression of CX26(Thr135Ala) did not affect neocortical excitatory neuron migration<sup>10</sup> (Supplementary Fig. 5), maturation (Supplementary Fig. 6) or chemical synapse formation in general (Supplementary Figs 7 and 8). Taken together, these results suggest that the blockade of electrical communication between sister excitatory neurons impairs the subsequent formation of chemical synapses between them.



**Figure 4 | Preferential electrical coupling is required for chemical synapse formation between sister excitatory neurons in neocortical ontogenetic columns.** a–f, Quadruple whole-cell recordings of two sister excitatory neurons (2 and 4) in ontogenetic columns expressing CX26–IRES–eGFP (a–c) or CX26(Thr135Ala)–IRES–eGFP (d–f) and of two adjacent non-sibling excitatory neurons (1 and 3). DIC and fluorescence images (a, d) and morphological reconstruction (b, e) of the respective quadruple whole-cell recordings are shown. The numbers 1 to 6 (left in b and e) indicate layers 1–6. Scale bars, 100  $\mu$ m (a, d). A summary of the chemical synaptic connections detected in the quadruple recordings is also shown (c, f). Pink shading indicates the existence of chemical synapses. Scale bars, 10 pA and 200 ms (c, f). AP, action potential; post, postsynaptic neuron; pre, presynaptic neuron. g, Summary of the frequency of chemical synapse formation between sister excitatory neurons in ontogenetic columns expressing eGFP, CX26–IRES–eGFP or CX26(Thr135Ala)–IRES–eGFP and their adjacent non-eGFP-expressing excitatory neurons. The bracket highlights the difference in the connectivity of CX26–IRES–eGFP-expressing sister neuron pairs and CX26(Thr135Ala)–IRES–eGFP-expressing sister neuron pairs.

Transient electrical coupling occurs before the establishment of mature patterns of synaptic connectivity in many developing nervous systems, and in some cases, this electrical coupling is crucial for the development of chemical synapses<sup>25–27</sup>. Previously, dye coupling was observed in the developing neocortex<sup>2,6,14,15</sup>. In addition, it has been suggested that the formation of electrically coupled neuronal domains might help to guide the emergence of chemically transmitting neuronal circuits<sup>7,8</sup>. However, electrically coupled neuronal pairs have not been reported in the neocortex at early developmental stages<sup>19</sup>. By performing dual and quadruple whole-cell recording experiments, which allow the detection of gap-junction-mediated electrical coupling with high sensitivity and spatial precision<sup>12,18</sup>, we demonstrated the electrical coupling of excitatory neurons in the early postnatal neocortex. Moreover, we revealed that neocortical excitatory neurons show a high preference for forming strong electrical coupling with their sister excitatory neurons but not with nearby non-sister excitatory neurons.



Furthermore, we found that strong electrical coupling between sister excitatory neurons in ontogenetic columns promotes their action potential generation and synchronous firing. Although previous studies have suggested that electrical coupling is crucial for robust synchronous activity in the neonatal neocortex<sup>8,28,29</sup>, the precise function of electrical coupling has been elusive. Our results show that electrical transmission between sister excitatory neurons in ontogenetic columns is required for the development of precise chemical synapses between these neurons. These findings provide clear evidence of the role of gap junctions in regulating precise neuronal circuit assembly in the neocortex.

## METHODS SUMMARY

Replication-incompetent Moloney murine leukaemia retroviruses expressing eGFP (obtained from F. H. Gage) or avian RCAS (replication-competent ASLV long terminal repeat with a splice acceptor) retroviruses expressing eGFP, CX26-IRES-eGFP or CX26(Thr135Ala)-IRES-eGFP were intraventricularly injected into E12–E13 CD-1 (Charles River Laboratories) or nestin-TVA-transgenic mouse embryos, respectively. Acute cortical slices were prepared at various postnatal stages, and multiple-electrode whole-cell recordings were performed on eGFP-expressing excitatory neurons in individual radial clones and on their excitatory neuron neighbours. Recordings were collected and analysed using two Axon Multiclamp 700B amplifiers and pCLAMP 10 (Molecular Devices) and IGOR 5 (WaveMetrics) software. Images were collected by confocal laser scanning microscopy (FluoView FV1000, Olympus) and analysed using FluoView (Olympus), NeuroLucida (MicroBrightField), Imaris (Andor Technology) and Photoshop (Adobe). Data are presented as mean  $\pm$  s.e.m., and statistical differences were determined using non-parametric statistical tests: the Mann–Whitney–Wilcoxon and Kruskal–Wallis tests.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 22 September 2011; accepted 14 February 2012.**

**Published online 2 May 2012.**

- Kriegstein, A. & Alvarez-Buylla, A. The glial nature of embryonic and adult neural stem cells. *Annu. Rev. Neurosci.* **32**, 149–184 (2009).
- Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714–720 (2001).
- Rakic, P. Specification of cerebral cortical areas. *Science* **241**, 170–176 (1988).
- Luskin, M. B., Pearlman, A. L. & Sanes, J. R. Cell lineage in the cerebral cortex of the mouse studied *in vivo* and *in vitro* with a recombinant retrovirus. *Neuron* **1**, 635–647 (1988).
- Yu, Y. C., Bultje, R. S., Wang, X. & Shi, S. H. Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature* **458**, 501–504 (2009).
- Peinado, A., Yuste, R. & Katz, L. C. Extensive dye coupling between rat neocortical neurons during the period of circuit formation. *Neuron* **10**, 103–114 (1993).
- Yuste, R., Peinado, A. & Katz, L. C. Neuronal domains in developing neocortex. *Science* **257**, 665–669 (1992).
- Yuste, R., Nelson, D. A., Rubin, W. W. & Katz, L. C. Neuronal domains in developing neocortex: mechanisms of coactivation. *Neuron* **14**, 7–17 (1995).
- Willecke, K. *et al.* Structural and functional diversity of connexin genes in the mouse and human genome. *Biol. Chem.* **383**, 725–737 (2002).
- Elias, L. A., Wang, D. D. & Kriegstein, A. R. Gap junction adhesion is necessary for radial migration in the neocortex. *Nature* **448**, 901–907 (2007).
- Nadarajah, B., Jones, A. M., Evans, W. H. & Parnavelas, J. G. Differential expression of connexins during neocortical development and neuronal circuit formation. *J. Neurosci.* **17**, 3096–3111 (1997).
- Connors, B. W. & Long, M. A. Electrical synapses in the mammalian brain. *Annu. Rev. Neurosci.* **27**, 393–418 (2004).
- Bennett, M. V. & Zukin, R. S. Electrical coupling and neuronal synchronization in the mammalian brain. *Neuron* **41**, 495–511 (2004).
- Lo Turco, J. J. & Kriegstein, A. R. Clusters of coupled neuroblasts in embryonic neocortex. *Science* **252**, 563–566 (1991).
- Connors, B. W., Benardo, L. S. & Prince, D. A. Coupling between neurons of the developing rat neocortex. *J. Neurosci.* **3**, 773–782 (1983).
- Gibson, J. R., Beierlein, M. & Connors, B. W. Two networks of electrically coupled inhibitory neurons in neocortex. *Nature* **402**, 75–79 (1999).
- Galarreta, M. & Hestrin, S. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature* **402**, 72–75 (1999).
- Wang, Y., Barakat, A. & Zhou, H. Electrotonic coupling between pyramidal neurons in the neocortex. *PLoS ONE* **5**, e10253 (2010).
- Pangratz-Fuehrer, S. & Hestrin, S. Synaptogenesis of electrical and GABAergic synapses of fast-spiking inhibitory neurons in the neocortex. *J. Neurosci.* **31**, 10767–10775 (2011).
- Stevens, C. F. & Zador, A. M. Input synchrony and the irregular firing of cortical neurons. *Nature Neurosci.* **1**, 210–217 (1998).
- Hebb, D. O. *The Organization of Behavior* (Wiley, 1949).
- Katz, L. C. & Shatz, C. J. Synaptic activity and the construction of cortical circuits. *Science* **274**, 1133–1138 (1996).
- Beahm, D. L. *et al.* Mutation of a conserved threonine in the third transmembrane helix of  $\alpha$ - and  $\beta$ -connexins creates a dominant-negative closed gap junction channel. *J. Biol. Chem.* **281**, 7994–8009 (2006).
- Rouan, F. *et al.* Trans-dominant inhibition of connexin-43 by mutant connexin-26: implications for dominant connexin disorders affecting epidermal differentiation. *J. Cell Sci.* **114**, 2105–2113 (2001).
- Personius, K. E. & Balice-Gordon, R. J. Loss of correlated motor neuron activity during synaptic competition at developing neuromuscular synapses. *Neuron* **31**, 395–408 (2001).
- Chuang, C. F., Vanhoven, M. K., Fetter, R. D., Verselis, V. K. & Bargmann, C. I. An innexin-dependent cell network establishes left–right neuronal asymmetry in *C. elegans*. *Cell* **129**, 787–799 (2007).
- Curtin, K. D., Zhang, Z. & Wyman, R. J. Gap junction proteins expressed during development are required for adult neural function in the *Drosophila* optic lamina. *J. Neurosci.* **22**, 7088–7096 (2002).
- Dupont, E., Hanganu, I. L., Kilb, W., Hirsch, S. & Luhmann, H. J. Rapid developmental switch in the mechanisms driving early cortical columnar networks. *Nature* **439**, 79–83 (2006).
- Kandler, K. & Katz, L. C. Coordination of neuronal activity in developing visual cortex by gap junction-mediated biochemical communication. *J. Neurosci.* **18**, 1419–1427 (1998).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the following: C. I. Bargmann, Y. Dan, A. L. Joyner, K. M. Hively and Y. Chin for comments on the manuscript; F. H. Gage for the 293gp NIT–GFP package cell line; E. C. Holland for nestin-TVA transgenic mice; and members of the Shi and Yu laboratories for their input. This work was supported by grants from the Ministry of Science and Technology of China (2012CB966300), the Natural Science Foundation of China (31121061 and 31070947), the Pujiang Talent Project of the Shanghai Science and Technology Committee (10PJ1400700), the Foundation of the Ministry of Education of China (20100071120061) (Y.-C.Y.), the National Institutes of Health (R01DA024681 and R21NS072483 (S.-H.S.), R21MH083624 (S.-H.S. and K.H.) and R01GM065947 (G.E.S.)), the McKnight Foundation and the March of Dimes Foundation (S.-H.S.).

**Author Contributions** Y.-C.Y., S.H. and S.-H.S. conceived the project. Y.-C.Y. and S.H. conducted the electrophysiology and morphology reconstruction experiments. S.C., Y.F. and K.N.B. generated viruses and performed *in utero* virus injections and morphological reconstructions. X.-H.Y., J.M. and K.P.G. performed immunohistochemistry and morphological reconstruction experiments. G.E.S. helped with CX26-carrying retrovirus engineering. K.H. performed cross-correlogram analysis. Y.-C.Y., S. H. and S.-H.S. analysed the data, interpreted the results and wrote the manuscript. All authors edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to Y.-C.Y. ([yccyu@fudan.edu.cn](mailto:yccyu@fudan.edu.cn)), S.H. ([hes@mshcc.org](mailto:hes@mshcc.org)) or S.-H.S. ([shis@mshcc.org](mailto:shis@mshcc.org)).

## METHODS

**Retrovirus production and *in utero* intraventricular injection.** Replication-incompetent eGFP-expressing Moloney murine leukaemia virus was produced from a stably transfected packaging cell line (293gp NIT-GFP, obtained from F. H. Gage) as previously reported<sup>2</sup>. Animals were maintained according to protocols approved by the Institutional Animal Care and Use Committee at the Sloan-Kettering Institute for Cancer Research and by Fudan University. *In utero* intraventricular injection was performed as previously described<sup>5</sup>. In brief, the uterine horns of pregnant CD-1 mice (Charles River Laboratories) at the E12–E13 stage of gestation were exposed in a clean environment. Retrovirus (~1.0  $\mu$ l) with Fast Green (2.5 mg ml<sup>-1</sup>, Sigma) was injected into the embryonic cerebral ventricle through a bevelled, calibrated glass micropipette (Drummond Scientific). After injection, the peritoneal cavity was lavaged with ~10 ml warm PBS (pH 7.4) containing antibiotics; the uterine horns were replaced; and the wound was closed. Avian RCAS (replication-competent ASLV long terminal repeat with a splice acceptor) retroviruses expressing CX26–IRES–eGFP or CX26(Thr135Ala)–IRES–eGFP were generated as previously described<sup>30</sup>. Similar *in utero* intraventricular injection of RCAS retroviruses was performed on the embryos of nestin-TVA-transgenic mice that were generated previously<sup>31</sup>.

**Immunohistochemistry and confocal imaging.** For CX26 and TUJ1 immunohistochemistry, after intracardial perfusion with cold PBS (pH 7.4) and 4% paraformaldehyde (PFA) in PBS (pH 7.4), the brains were removed and embedded in the tissue freezing medium OCT compound (Electron Microscopy Sciences) after sucrose treatment. Coronal sections (20  $\mu$ m) were prepared using a cryostat (Leica Microsystems) and incubated for 2 h at room temperature in a blocking solution (10% normal goat serum and 0.1% Triton X-100 in PBS), followed by incubation with primary antibodies, including rabbit anti-CX26 antibody (Invitrogen) and mouse anti-TUJ1 antibody (Covance) for 2 days at 4 °C. Sections were then washed in 0.1% Triton X-100 in PBS and incubated with the appropriate secondary antibodies overnight at 4 °C. For three-dimensional reconstruction, Z-series images were taken at 0.1- $\mu$ m steps using a 100 $\times$  objective lens in an Olympus FluoView FV1000 confocal laser scanning microscope and were analysed using Neurolucida (MicroBrightField) and Imaris (Andor Technology).

**Slice preparation, electrophysiological recording and data analysis.** Embryos that received retroviral injections were delivered naturally. Brains were removed at various times after birth, and acute cortical slices (200–300  $\mu$ m thick) were prepared in artificial cerebrospinal fluid (ACSF) containing 126 mM NaCl, 3 mM KCl, 1.25 mM KH<sub>2</sub>PO<sub>4</sub>, 1.3 mM MgSO<sub>4</sub>, 3.2 mM CaCl<sub>2</sub>, 26 mM NaHCO<sub>3</sub> and 10 mM glucose, bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>, with a vibratome (Leica Microsystems) at 4 °C. Slices were allowed to recover in an interface chamber at 35 °C for at least 1 h and were then kept at room temperature before being transferred to a recording chamber containing ACSF at 34 °C. An infrared DIC microscope (Olympus) equipped with epifluorescence illumination, a charge-coupled device camera and two water immersion lenses (10 $\times$  and 60 $\times$ ) was used to visualize and target the recording electrodes to eGFP-expressing sister cells in ontogenetic columns and to their nearby control cells. Glass recording electrodes (20–30 M $\Omega$  resistance) were filled with an intracellular solution consisting of 130 mM potassium gluconate, 6 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM EGTA, 10 mM HEPES, 2.5 mM Na<sub>2</sub>ATP, 0.5 mM Na<sub>2</sub>GTP, 10 mM potassium phosphocreatine and 0.3% Alexa Fluor 568 hydrazide (Invitrogen) (pH 7.25 and 295 mOsmol per kg solution). Recordings were collected and analysed using an Axon Multiclamp 700B amplifier and pCLAMP 10 software (Molecular Devices). In all dual and quadruple recordings, electrical coupling was assessed by injecting currents to trigger hyperpolarization or depolarization under current-clamp mode or by

voltage steps under voltage-clamp mode. In some experiments, 100  $\mu$ M MFA (Sigma) was added to the bath to block gap junctions, and 10  $\mu$ M BMI, 20  $\mu$ M CPP and 10  $\mu$ M CNQX (Tocris Bioscience) were used to block GABA<sub>A</sub>, NMDA and AMPA receptors, respectively. MFA was applied for 10–30 min to block gap junctions and was later washed out for 20–30 min to test the recovery of the gap junctions. The gap junction electrical conductance was estimated under voltage-clamp conditions<sup>32</sup>. Native neuronal activity stimuli were simulated as previously described<sup>20</sup>.

Chemical connections between neuron pairs were assessed by injecting current to induce action potentials in one of the neurons kept in current clamp while testing for postsynaptic responses in other neurons under the voltage-clamp recording condition at –70 mV. Glass recording electrodes (8–12 M $\Omega$  resistance) were filled with an intracellular solution consisting of 130 mM potassium gluconate, 6 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM EGTA, 10 mM HEPES, 2.5 mM Na<sub>2</sub>ATP, 0.5 mM Na<sub>2</sub>GTP, 10 mM potassium phosphocreatine and 0.5% neurobiotin. For every possible pair, the connections were tested in both directions for at least 20 trials, with both single action potentials and trains of action potentials being generated in each presynaptic neuron. For miniature excitatory postsynaptic current (mEPSC) and miniature inhibitory postsynaptic current (mIPSC) analysis, cells were clamped at –70 mV, and recordings were performed in the presence of 50  $\mu$ M BMI or 10  $\mu$ M CNQX/20  $\mu$ M D-AP5 together with 1  $\mu$ M tetrodotoxin, respectively. The decay time constant,  $\tau$ , was estimated by single-exponential function fitting:  $f(t) = A \cdot \exp(-t/\tau) + C$ , where  $A$  and  $C$  are constants and  $t$  is time. In whole-cell patch-clamp recording experiments, slices were fixed in 4% PFA in PBS (pH 7.4) after the recordings were completed, and the morphology of recorded neurons that had been loaded with Alexa Fluor 568 hydrazide through the recording pipette was visualized using an Olympus FluoView FV1000 confocal laser scanning microscope. Z-series images were taken at 1–3- $\mu$ m steps and analysed using FluoView (Olympus), Neurolucida (MicroBrightField) and Photoshop (Adobe). In some recording experiments, neurobiotin was later visualized with Alexa-Fluor-647- or Alexa-Fluor 568-conjugated streptavidin (Invitrogen).

Normalized cross-correlograms of firing patterns were analysed as previously described<sup>33</sup>. In brief, the number of times that neuron 1 fired within a time interval ( $n\Delta t$ ,  $(n + 1)\Delta t$ ) from spikes fired by neuron 2 was calculated (and is denoted  $y_n$ , which is the number of counts per bin, where the bin width is  $\Delta t = 1$  ms). The cross-correlogram  $y_n$  was normalized to standard scores:  $Z = (y_n - \gamma_E)/s_y$ , where  $\gamma_E = f_1 \cdot f_2 \cdot T \cdot \Delta t$ , and  $f_{1,2}$  is the average firing rate of neurons 1 and 2,  $T$  is the recording time and  $s_y$  is the standard deviation of  $y_n$ . Data are presented as mean  $\pm$  s.e.m., and statistical differences were determined using non-parametric statistical tests: the Mann–Whitney–Wilcoxon and Kruskal–Wallis tests. Peaks in the cross-correlogram were considered significant if individual bins exceeded the expected value by three standard deviations (that is, if the  $Z$  score was  $>3$ ).

30. Du, Z. *et al.* Introduction of oncogenes into mammary glands *in vivo* with an avian retroviral vector initiates and promotes carcinogenesis in mouse models. *Proc. Natl Acad. Sci. USA* **103**, 17396–17401 (2006).
31. Holland, E. C., Hively, W. P., DePinho, R. A. & Varmus, H. E. A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. *Genes Dev.* **12**, 3675–3685 (1998).
32. Neyton, J. & Trautmann, A. Single-channel currents of an intercellular junction. *Nature* **317**, 331–335 (1985).
33. Vos, B. P., Maex, R., Volny-Luraghi, A. & De Schutter, E. Parallel fibers synchronize spontaneous activity in cerebellar Golgi cells. *J. Neurosci.* **19**, RC6 (1999).

# An ultraviolet–optical flare from the tidal disruption of a helium–rich stellar core

S. Gezari<sup>1</sup>, R. Chornock<sup>2</sup>, A. Rest<sup>3</sup>, M. E. Huber<sup>4</sup>, K. Forster<sup>5</sup>, E. Berger<sup>2</sup>, P. J. Challis<sup>2</sup>, J. D. Neill<sup>5</sup>, D. C. Martin<sup>5</sup>, T. Heckman<sup>1</sup>, A. Lawrence<sup>6</sup>, C. Norman<sup>1</sup>, G. Narayan<sup>2</sup>, R. J. Foley<sup>2</sup>, G. H. Marion<sup>2</sup>, D. Scolnic<sup>1</sup>, L. Chomiuk<sup>2</sup>, A. Soderberg<sup>2</sup>, K. Smith<sup>7</sup>, R. P. Kirshner<sup>2</sup>, A. G. Riess<sup>1</sup>, S. J. Smartt<sup>7</sup>, C. W. Stubbs<sup>2</sup>, J. L. Tonry<sup>4</sup>, W. M. Wood-Vasey<sup>8</sup>, W. S. Burgett<sup>4</sup>, K. C. Chambers<sup>4</sup>, T. Grav<sup>9</sup>, J. N. Heasley<sup>4</sup>, N. Kaiser<sup>4</sup>, R.-P. Kudritzki<sup>4</sup>, E. A. Magnier<sup>4</sup>, J. S. Morgan<sup>4</sup> & P. A. Price<sup>10</sup>

**The flare of radiation from the tidal disruption and accretion of a star can be used as a marker for supermassive black holes that otherwise lie dormant and undetected in the centres of distant galaxies<sup>1</sup>. Previous candidate flares<sup>2–6</sup> have had declining light curves in good agreement with expectations, but with poor constraints on the time of disruption and the type of star disrupted, because the rising emission was not observed. Recently, two ‘relativistic’ candidate tidal disruption events were discovered, each of whose extreme X-ray luminosity and synchrotron radio emission were interpreted as the onset of emission from a relativistic jet<sup>7–10</sup>. Here we report a luminous ultraviolet–optical flare from the nuclear region of an inactive galaxy at a redshift of 0.1696. The observed continuum is cooler than expected for a simple accreting debris disk, but the well-sampled rise and decay of the light curve follow the predicted mass accretion rate and can be modelled to determine the time of disruption to an accuracy of two days. The black hole has a mass of about two million solar masses, modulo a factor dependent on the mass and radius of the star disrupted. On the basis of the spectroscopic signature of ionized helium from the unbound debris, we determine that the disrupted star was a helium-rich stellar core.**

When the pericentre of a star’s orbit ( $R_p$ ) passes within the tidal disruption radius of a massive black hole,  $R_T \approx R_*(M_{BH}/M_*)^{1/3}$  (where  $R_*$  is the stellar radius,  $M_{BH}$  is the black-hole mass and  $M_*$  is the stellar mass), tidal forces overcome the binding energy of the star, which breaks up with roughly half of the stellar debris remaining bound to the black hole and the rest being ejected at high velocity<sup>1</sup>. For black holes above a critical mass,  $M_{crit} \approx 10^8 r_*^{3/2} m_*^{-1/2} M_\odot$  (where  $r_* = R_*/R_\odot$ ,  $m_* = M_*/M_\odot$ ,  $R_\odot$  is the solar radius and  $M_\odot$  is the solar mass), the star becomes trapped within the event horizon of the black hole before being disrupted. The mass accretion rate ( $\dot{M}$ ) in a tidal disruption event can be calculated directly from the orbital return times of the bound debris<sup>11,12</sup>. For the simplest case, of a star of uniform density, this yields  $\dot{M} = (2/3)(fM_*/t_{min})(t/t_{min})^{-5/3}$ , where  $f$  is the fraction of the star accreted and  $t_{min}$  is the orbital period of the most tightly bound debris and, therefore, the time delay between the time of disruption and the start of the flare, which scales as  $M_{BH}^{1/2} M_*^{-1} R_*^{3/2}$  for  $R_p = R_T$ . The radiative output of the accreted debris is less certain, and depends on the ratio of the accretion rate to the Eddington rate<sup>13</sup>.

The optical transient, PS1-10jh (right ascension,  $\alpha_{J2000} = 16^h 09^m 28.296^s$ ; declination,  $\delta_{J2000} = +53^\circ 40' 23.52''$ ), was discovered on 2010 May 31.45 UT (universal time) in the Pan-STARRS1<sup>14</sup> (PS1) Medium Deep Survey by our two independent image-differencing pipelines. The densely sampled (cadence,  $\Delta t \approx 3$  d) optical light curves of PS1-10jh in the  $g_{P1}$ ,  $r_{P1}$ ,  $i_{P1}$  and  $z_{P1}$  bands (Supplementary Information) follow the rise of the transient to its peak in the  $g_{P1}$  band

on 2010 July 12.31 UT and its subsequent decay until 2011 September 1.24 UT (Supplementary Table 1). PS1-10jh was discovered independently as a transient, near-ultraviolet (NUV) source at the  $20\sigma$  level by the Galaxy Evolution Explorer<sup>15</sup> (GALEX) Time Domain Survey (TDS) on 2010 June 17.68 UT within  $2.5 \pm 3.0$  arcsec of the PS1 location, and was detected in ten more epochs of TDS observations between then and 2011 June 10.68 UT (Supplementary Table 2). No source is detected in a coaddition of all the TDS epochs in 2009, with a  $3\sigma$  upper limit of  $>25.6$  mag implying a peak amplitude of variability in the NUV of  $>6.4$  mag. See Supplementary Information for details on the PS1 and GALEX photometry.

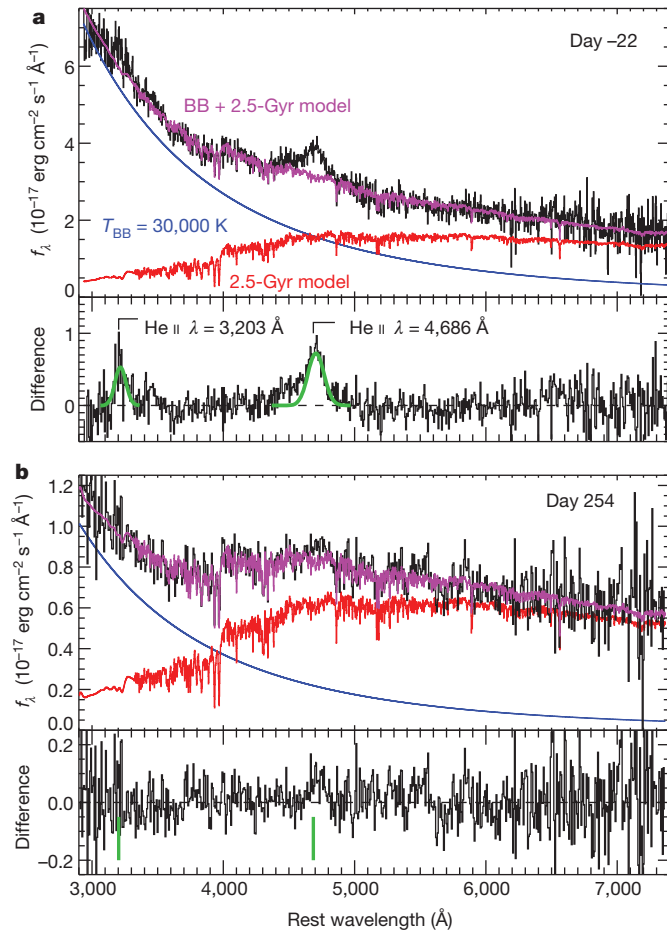
PS1-10jh is coincident with the centre of a galaxy within the  $3\sigma$  positional uncertainty (0.036 arcsec; Supplementary Information), with rest-frame u-, g-, r-, i- and z-band photometry from the Sloan Digital Sky Survey<sup>16</sup> and K-band photometry from the UK Infrared Telescope Infrared Sky Survey<sup>17</sup> fitted with a galaxy template<sup>18</sup> with  $M_{stars} = (3.6 \pm 0.2) \times 10^9 M_\odot$  and  $M_r = -18.7$  mag, where  $M_{stars}$  is the galaxy stellar mass and  $M_r$  is the absolute r-band magnitude. The mass of the central black hole as determined indirectly from locally established scaling relations<sup>19</sup> is  $4^{+4}_{-2} \times 10^6 M_\odot$ . We obtained five epochs of optical spectroscopy at the location of PS1-10jh between 2010 June 16.33 and 2011 September 4.23 UT with the 6.5-m MMT (Supplementary Table 3). The continua in the spectra are well modelled by the combination of a galaxy host at redshift  $z = 0.1696$  (luminosity distance, 816 Mpc) with a stellar population with an age of 1.4–5.0 Gyr, depending on the chosen metallicity, and a fading hot blackbody component with temperature  $T_{BB} \approx 3 \times 10^4$  K (Fig. 1).

The spectra show no narrow emission lines that would be indicative of star formation or an active galactic nucleus (AGN). We obtained a 10-ks, 0.2–10-keV X-ray observation, using the Chandra X-ray Observatory, at the location of PS1-10jh on 2011 May 22.96 UT, and detected no source above the background with a  $3\sigma$  upper limit of  $L_X(0.2–10 \text{ keV}) < 5.8 \times 10^{41} \text{ erg s}^{-1}$  for an unobscured AGN spectrum. The X-ray faintness and extreme NUV variability amplitude of PS1-10jh strongly disfavour its origin in an AGN, and its prolonged brightness in the ultraviolet strongly disfavours its origin in a supernova (Supplementary Information).

The rise and decay of the light curve of PS1-10jh is well described by numerical simulations of the mass return rate from a star that is tidally disrupted at  $R_p = R_T$  and has an internal structure parameterized by a polytropic exponent of 5/3 characteristic of a fully convective star or a degenerate core<sup>20</sup> (Fig. 2). The decay from the peak is too steep to be fitted by simulations of a more centrally concentrated stellar structure, such as one that is characteristic of a solar-type star (Supplementary Information). There are systematic differences between the light curve

<sup>1</sup>Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA. <sup>2</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA. <sup>3</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. <sup>4</sup>Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA. <sup>5</sup>California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. <sup>6</sup>Institute for Astronomy, University of Edinburgh Scottish Universities Physics Alliance, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK. <sup>7</sup>Astrophysics Research Centre, School of Mathematics and Physics, Queen’s University Belfast, Belfast BT7 1NN, UK. <sup>8</sup>Pittsburgh Particle Physics, Astrophysics, and Cosmology Center, Department of Physics and Astronomy, University of Pittsburgh, 3941 O’Hara Street, Pittsburgh, Pennsylvania 15260, USA. <sup>9</sup>Planetary Science Institute, 1700 East Fort Lowell, Tucson, Arizona 85719, USA. <sup>10</sup>Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA.

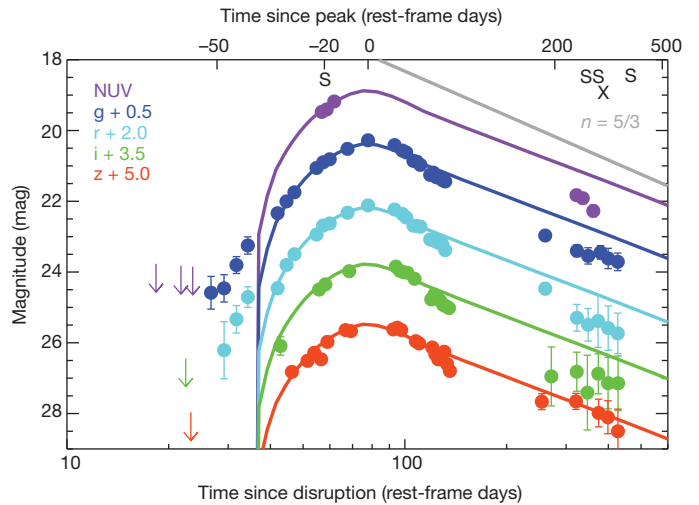




**Figure 1 | Optical spectrum.** MMT optical spectra (black) of PS1-10jh obtained  $-22$  (a) and  $254$  (b) rest-frame days from the peak, expressed in terms of flux density. Each continuum is fitted with a combination (magenta) of a stellar population  $2.5$  Gyr old (red) and a fading blackbody with a temperature of  $\sim 3 \times 10^4$  K determined from the ultraviolet–optical spectral energy distribution (SED) (blue). The difference between the black and magenta spectra is shown in the lower part of each panel. Helium II emission at  $\lambda = 4,686$  Å (Fowler series,  $n = 4 \rightarrow 3$ ) is detected above the continuum model and fitted with a Gaussian with a full-width at half-maximum of  $9,000 \pm 700$  km s $^{-1}$  and luminosity  $L = (9 \pm 1) \times 10^{40}$  erg s $^{-1}$  (plotted with a green line in the early epoch (a)). Residual emission above the continuum model is also detected at  $\sim 3,200$  Å, which is coincident with the location of the He II  $\lambda = 3,203$  Å (Fowler series,  $n = 5 \rightarrow 3$ ) line, and confirms the identification of He II  $\lambda = 4,686$  Å emission. The observed flux ratio of He II  $\lambda = 3,203$  Å emission to He II  $\lambda = 4,686$  Å emission is  $0.50 \pm 0.10$ , measured using a Gaussian fit to the  $\lambda = 3,203$  Å line with a width fixed to that of the  $\lambda = 4,686$  Å line, limits the internal extinction to  $E(B - V) < 0.08$  mag (Supplementary Information). The He II  $\lambda = 4,686$  Å line is still evident as an excess above the model in the later epoch (b), but it has faded by a factor of  $\sim 10$  since 22 rest-frame days before the peak, the same factor by which the ultraviolet continuum has faded during this time. The absolute flux scaling in the later epoch is uncertain owing to obscuration by clouds on the date of the observation.

and the model during the early rise (more than 44 rest-frame days before the peak) and the late decay (more than 240 rest-frame days after the peak), which could imply a stellar structure more complex than one described by a single polytrope. The mass of the black hole is determined from the stretch factor of  $1.38 \pm 0.03$  applied to fit the model of a  $10^6 M_{\odot}$  black hole to the light curve, which implies that the time of disruption was  $76 \pm 2$  d before the peak and that  $M_{\text{BH}} = (1.9 \pm 0.1) \times 10^6 m_{*}^2 r_{*}^{-3} M_{\odot}$ .

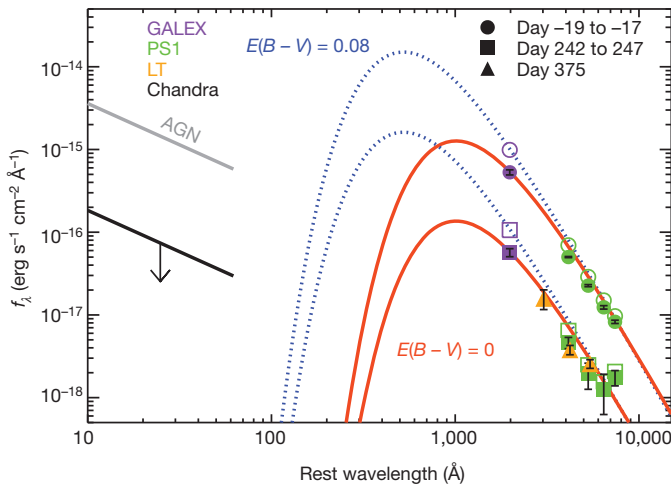
The most constraining property of PS1-10jh is the detection of very broad high-ionization He II emission at wavelengths of  $\lambda = 4,686$  Å



**Figure 2 | Ultraviolet–optical light curve.** The GALEX NUV and PS1  $g_{\text{P1}}$ ,  $r_{\text{P1}}$ ,  $i_{\text{P1}}$  and  $z_{\text{P1}}$ -band light curves of PS1-10jh (with the flux from the host galaxy removed), plotted against logarithmic time since the peak (top axis) and since the disruption (bottom axis). The curves (shown with solid lines scaled to the flux in the GALEX and PS1 bands) were determined from the best fit of the  $g_{\text{P1}}$ -band light curve to a numerical model<sup>20</sup> of the mass accretion rate of a tidally disrupted star with a polytropic exponent of  $5/3$ . For each of the four optical bands, we independently performed a least-squares fit of the model for a  $10^6 M_{\odot}$  black hole to the light curve from  $-36$  to  $58$  rest-frame days from the peak, with the time of disruption, a vertical scaling factor and a time stretch factor as free parameters. The GALEX and PS1 photometry at  $t > 240$  rest-frame days since the peak is shown binned in time to increase the signal-to-noise ratio. The dates of multiple epochs of MMT spectroscopy are marked with an S, and the date of the Chandra X-ray observation is marked with an X. The grey line shows an  $n = 5/3$  power-law decay from the peak. Errors,  $1\sigma$ ; arrows,  $3\sigma$  upper limits.

(full-width at half-maximum,  $9,000 \pm 700$  km s $^{-1}$ ) and  $3,203$  Å that fade in time along with the ultraviolet–optical continuum. The lack of Balmer line emission in the spectra requires an extremely low hydrogen mass fraction, of  $< 0.2$  (Supplementary Information), which cannot be found in the ambient interstellar medium or in a passive accretion disk. This is the strongest evidence that PS1-10jh must be fuelled by the accretion of a star that has lost its hydrogen envelope, either through stellar winds or through tidal interactions with the central supermassive black hole. The broad width of the line is also what is expected from the velocities of the most energetic unbound stellar debris in a tidal disruption event<sup>21</sup>, that is  $v_{\text{max}} \approx 1 \times 10^4 (M_{\text{BH}}/10^6 M_{\odot})^{1/6} (R_{\text{T}}/R_{\text{P}})^{-1/2} m_{*}^{-1/3}$  km s $^{-1}$ .

We measure the SED of the flare over time from the nearly simultaneous PS1 optical and GALEX ultraviolet observations (with the host galaxy flux removed; Fig. 3). The pre-peak SED is fitted by a blackbody with  $T_{\text{BB}} = (2.9 \pm 0.2) \times 10^4$  K, consistent with the blackbody component seen in the spectra. However, the temperature fit is very sensitive to internal extinction. If we correct for the maximum internal extinction of  $E(B - V) = 0.08$  mag allowed by the ratio between the observed He II  $\lambda = 3,203$  Å and  $\lambda = 4,686$  Å emission, the best-fit temperature increases to  $(5.5 \pm 0.4) \times 10^4$  K. In fact, we know that the photo-ionizing continuum must have  $T_{\text{BB}} \gtrsim 5 \times 10^4$  K 22 rest-frame days before the peak to produce enough  $\lambda < 228$  Å photons to photo-ionize the He II  $\lambda = 4,686$  Å line observed with a luminosity of  $(9 \pm 1) \times 10^{40}$  erg s $^{-1}$ . The late-time SED can be fitted with the same temperature as the pre-peak SED. We note that the observed continuum temperature, and even the maximum temperature allowed by possible de-reddening, is considerably cooler than the temperature of  $\sim 2.5 \times 10^5 (M_{\text{BH}}/10^6 M_{\odot})^{1/12} r_{*}^{-1/2} m_{*}^{-1/6}$  K expected from material radiating at the Eddington limit at the tidal radius<sup>13</sup>. This discrepancy is also seen in AGNs<sup>22</sup> and may imply that the continuum we see is due to reprocessing of some kind<sup>22,23</sup>.



**Figure 3 | Spectral energy distribution.** SED of PS1-10jh during nearly simultaneous GALEX ultraviolet and PS1 optical observations (with the flux from the host galaxy removed) at two epochs (rest-frame days  $-19$  to  $-17$  and  $242$  to  $247$  from the peak of the flare). Flux densities have been corrected for galactic extinction of  $E(B-V) = 0.013$  mag. The ultraviolet-optical SED from  $-19$  to  $247$  rest-frame days from the peak is fitted with a  $2.9 \times 10^4$  K blackbody. Orange solid lines show blackbodies with this temperature scaled to the respective NUV flux densities for the respective epochs. Open symbols show the GALEX and PS1 flux densities corrected for an internal extinction of  $E(B-V) = 0.08$  mag, and the dotted blue lines shows the  $5.5 \times 10^4$  K blackbody fit to the de-reddened flux densities for the respective epochs. The upper limit from the Chandra observation on 2011 May 22.96 UT assuming a spectrum with a photon index of  $\Gamma = 2$ , typical of an AGN, is plotted with a thick black line. The X-ray flux density expected from an unobscured AGN with a comparable NUV flux is plotted for comparison with a thick grey line<sup>28</sup>. Also shown are the u-, g- and r-band flux densities measured from aperture photometry with the Liverpool Telescope<sup>29</sup> (LT) on 2011 September 24.91 UT, after subtracting the flux from the host galaxy as measured by the Sloan Digital Sky Survey. Errors,  $1\sigma$ .

On the basis of the arguments above, we assume that the observed temperature is a lower limit,  $T_{\text{BB}} \gtrsim 3 \times 10^4$  K. The peak bolometric luminosity is thus  $\gtrsim 2.2 \times 10^{44}$  erg s $^{-1}$  and the total energy emitted from integrating under the light-curve model is  $\gtrsim 2.1 \times 10^{51}$  erg, corresponding to a total accreted mass ( $M_{\text{acc}}$ ) of  $\gtrsim 0.012(\epsilon/0.1)^{-1} M_{\odot}$ , where  $\epsilon$  is the efficiency of converting matter into radiation.

The internal structure and high helium abundance of the star derived from the light curve and the spectra can be consistently modelled as the tidally stripped core of a red giant (the precursor to a helium white dwarf) that had a main-sequence mass of  $M_{\star} \gtrsim 1 M_{\odot}$  so as to have evolved off the main sequence in less time than the age of the stellar population ( $< 5$  Gyr). This tidal stripping mechanism has been invoked to explain the hot stars in the Galactic Centre<sup>24</sup>, and the rate of tidal disruption of tidally stripped stars is likely to be higher than that of solar-type stars<sup>25</sup>. The mass of the black hole derived from the light-curve fit depends on the mass and radius of the star at the time of disruption. Using  $M_{\star} \approx 0.23 M_{\odot}$  and  $R_{\star} \approx 0.33 R_{\odot}$  (values measured for a red giant core that was stripped in a binary system<sup>26</sup>), and assuming that the evolution of the core is similar to one that is tidally stripped, we find that  $f = M_{\text{acc}}/M_{\star} \gtrsim 0.058$  (approaching  $f \gtrsim 0.1$  as measured in simulations<sup>27</sup>), that  $M_{\text{BH}} = (2.8 \pm 0.1) \times 10^6 M_{\odot}$  and that the peak luminosity approaches the Eddington luminosity of the supermassive black hole ( $L_{\text{peak}} \gtrsim 0.6 L_{\text{Edd}}$ ).

Received 8 November 2011; accepted 23 February 2012.

Published online 2 May 2012.

- Rees, M. J. Tidal disruption of stars by black holes of  $10$  to the  $6$ th- $10$  to the  $8$ th solar masses in nearby galaxies. *Nature* **333**, 523–528 (1988).
- Komossa, S. & Bade, N. The giant X-ray outbursts in NGC 5905 and IC 3599: follow-up observations and outburst scenarios. *Astron. Astrophys.* **343**, 775–787 (1999).

- Komossa, S. *et al.* A huge drop in the X-ray luminosity of the nonactive galaxy RX J1242.6–1119A, and the first postflare spectrum: testing the tidal disruption scenario. *Astrophys. J.* **603**, L17–L20 (2004).
- Esquej, P. *et al.* Evolution of tidal disruption candidates discovered by XMM-Newton. *Astron. Astrophys.* **489**, 543–554 (2008).
- Gezari, S. *et al.* Luminous thermal flares from quiescent supermassive black holes. *Astrophys. J.* **698**, 1367–1379 (2009).
- van Velzen, S. *et al.* Optical discovery of probable stellar tidal disruption flares. *Astrophys. J.* **741**, 73–96 (2011).
- Bloom, J. S. *et al.* A possible relativistic jetted outburst from a massive black hole fed by a tidally disrupted star. *Science* **333**, 203–206 (2011).
- Burrows, D. N. *et al.* Relativistic jet activity from the tidal disruption of a star by a massive black hole. *Nature* **476**, 421–424 (2011).
- Zauderer, B. A. *et al.* Birth of a relativistic outflow in the unusual  $\gamma$ -ray transient Swift J164449.3+573451. *Nature* **476**, 425–428 (2011).
- Cenko, S. B. *et al.* Swift J2058.4+0516: discovery of a possible second relativistic tidal disruption flare. Preprint at (<http://arxiv.org/abs/1107.5307>) (2011).
- Phinney, E. S. in *The Center of the Galaxy* (ed. Morris, M.) 543–553 (IAU Symp. 136, Kluwer, 1989).
- Evans, C. R. & Kochanek, C. S. The tidal disruption of a star by a massive black hole. *Astrophys. J.* **346**, L13–L16 (1989).
- Ulmer, A. Flares from the tidal disruption of stars by massive black holes. *Astrophys. J.* **514**, 180–187 (1999).
- Kaiser, N. *et al.* The Pan-STARRS wide-field optical/NIR imaging survey. *Proc. SPIE* **7733**, 77330E (2010).
- Martin, D. C. *et al.* The Galaxy Evolution Explorer: a space ultraviolet survey mission. *Astrophys. J.* **619**, L1–L6 (2005).
- Aihara, H. *et al.* The eighth data release of the Sloan Digital Sky Survey: first data from SDSS-III. *Astrophys. J. Suppl. Ser.* **193**, 29–45 (2011).
- Lawrence, A. *et al.* The UKIRT Infrared Deep Sky Survey (UKIDSS). *Mon. Not. R. Astron. Soc.* **379**, 1599–1617 (2007).
- Blanton, M. R. & Roweis, S. K-corrections and filter transformations in the ultraviolet, optical, and near-infrared. *Astron. J.* **133**, 734–754 (2007).
- Häring, N. & Rix, H.-W. On the black hole mass-bulge mass relation. *Astrophys. J.* **604**, L89–L92 (2004).
- Lodato, G., King, A. R. & Pringle, J. E. Stellar disruption by a supermassive black hole: is the light curve really proportional to  $t^{-5/3}$ ? *Mon. Not. R. Astron. Soc.* **392**, 332–340 (2009).
- Strubbe, L. E. & Quataert, E. Optical flares from the tidal disruption of stars by massive black holes. *Mon. Not. R. Astron. Soc.* **400**, 2070–2084 (2009).
- Lawrence, A. The UV peak in active galactic nuclei: a false continuum from blurred reflection? Preprint (<http://arxiv.org/abs/1110.0854>) (2011).
- Loeb, A. & Ulmer, A. Optical appearance of the debris of a star disrupted by a massive black hole. *Astrophys. J.* **489**, 573–578 (1997).
- Davies, M. B. & King, A. The stars of the galactic center. *Astrophys. J.* **624**, L25–L27 (2005).
- Kobayashi, S., Laguna, P., Phinney, E. S. & Mészáros, P. Gravitational waves and X-ray signals from stellar disruption by a massive black hole. *Astrophys. J.* **615**, 855–865 (2004).
- Moxed, P. F. L. *et al.* Discovery of a stripped red giant core in a bright eclipsing binary system. *Mon. Not. R. Astron. Soc.* **418**, 1156–1164 (2011).
- Ayal, S., Livio, M. & Piran, T. Tidal disruption of a solar-type star by a supermassive black hole. *Astrophys. J.* **545**, 772–780 (2000).
- Steffen, A. T. *et al.* The X-ray-to-optical properties of optically selected active galaxies over wide luminosity and redshift ranges. *Astron. J.* **131**, 2826–2842 (2006).
- Steele, I. A. *et al.* The Liverpool Telescope: performance and first results. *Proc. SPIE* **5389**, 679 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank H. Tananbaum for approving our Chandra Director's Discretionary Time request. We are grateful to G. Lodato for providing the tidal disruption event models in tabular form, and to S. Moran for running software to calculate the host-galaxy K-corrections. We thank R. E. Williams for discussions on the line emission in the spectra. S.G. was supported by NASA through a Hubble Fellowship grant awarded by the Space Telescope Science Institute, which is operated by AURA Inc. for NASA. Partial support for this work was provided by the National Science Foundation. The PS1 survey has been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Inc. and the National Central University of Taiwan, and by NASA under a grant issued through the Planetary Science Division of the NASA Science Mission Directorate. We acknowledge NASA's support for construction, operation, and science analysis of the GALEX mission, which was developed in cooperation with Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology. Some of the observations reported here were obtained at the MMT Observatory, which is a joint facility of the Smithsonian Institution and the University of Arizona, and at the Liverpool Telescope, which is operated with financial support from the UK Science and Technology Facilities Council. The computations in this paper were run on the Odyssey cluster supported by the FAS Science Division Research Computing Group at Harvard University. R.J.F. is a Clay Fellow.

**Author Contributions** S.G. designed the observations and the transient detection pipeline for the GALEX TDS, and measured the ultraviolet photometry of PS1-10jh. K.F.

and J.D.N coordinated, and D.C.M. facilitated, the GALEX TDS observations. A.R. designed the PhotPipe transient detection pipeline hosted by Harvard/CfA for the PS1 Medium Deep Survey (MDS), and measured the optical photometry of PS1-10jh. R.C. designed, implemented and analysed the MMT optical spectroscopy observations, and contributed to the operation of PhotPipe and the visual inspection of transient alerts. E.B. proposed and facilitated the MMT observations. M.E.H., G.N., D.S. and R.J.F. contributed to the operation of PhotPipe and the visual inspection of transient alerts. P.J.C., R.J.F., G.H.M., L.C. and A.S. contributed to the MMT observations. S.J.S. designed, and K.S. operated, the transient pipeline for PS1 MDS hosted by Queen's University Belfast. C.W.S., J.L.T. and W.M.W.-V. facilitated the transient pipelines for PS1 MDS. W.S.B., K.C.C., T.G., J.N.H., N.K., R.-P.K., E.A.M., J.S.M., P.A.P., C.W.S. and J.L.T. helped build the PS1 system. S.G. requested the Director's Discretionary Time Chandra X-ray

observation and analysed the data. A.L. obtained the Liverpool Telescope optical imaging observations and analysed the data, and stimulated discussions on the nature of the SED of PS1-10jh. S.G. analysed and modelled the multicolour light curve and the SED of PS1-10jh. T.H. and C.N. stimulated discussions on the nature of the disrupted star. The paper was organized and written by S.G., and all authors provided feedback on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.G. ([suvi@pha.jhu.edu](mailto:suvi@pha.jhu.edu)).



# Warming experiments underpredict plant phenological responses to climate change

E. M. Wolkovich<sup>1</sup>, B. I. Cook<sup>2,3</sup>, J. M. Allen<sup>4</sup>, T. M. Crimmins<sup>5</sup>, J. L. Betancourt<sup>6</sup>, S. E. Travers<sup>7</sup>, S. Pau<sup>8</sup>, J. Regetz<sup>8</sup>, T. J. Davies<sup>9</sup>, N. J. B. Kraft<sup>10,11</sup>, T. R. Ault<sup>12</sup>, K. Bolmgren<sup>13,14</sup>, S. J. Mazer<sup>15</sup>, G. J. McCabe<sup>16</sup>, B. J. McGill<sup>17</sup>, C. Parmesan<sup>18,19</sup>, N. Salamin<sup>20,21</sup>, M. D. Schwartz<sup>22</sup> & E. E. Cleland<sup>1</sup>

Warming experiments are increasingly relied on to estimate plant responses to global climate change<sup>1,2</sup>. For experiments to provide meaningful predictions of future responses, they should reflect the empirical record of responses to temperature variability and recent warming, including advances in the timing of flowering and leafing<sup>3–5</sup>. We compared phenology (the timing of recurring life history events) in observational studies and warming experiments spanning four continents and 1,634 plant species using a common measure of temperature sensitivity (change in days per degree Celsius). We show that warming experiments underpredict advances in the timing of flowering and leafing by 8.5-fold and 4.0-fold, respectively, compared with long-term observations. For species that were common to both study types, the experimental results did not match the observational data in sign or magnitude. The observational data also showed that species that flower earliest in the spring have the highest temperature sensitivities, but this trend was not reflected in the experimental data. These significant mismatches seem to be unrelated to the study length or to the degree of manipulated warming in experiments. The discrepancy between experiments and observations, however, could arise from complex interactions among multiple drivers in the observational data, or it could arise from remediable artefacts in the experiments that result in lower irradiance and drier soils, thus dampening the phenological responses to manipulated warming. Our results introduce uncertainty into ecosystem models that are informed solely by experiments and suggest that responses to climate change that are predicted using such models should be re-evaluated.

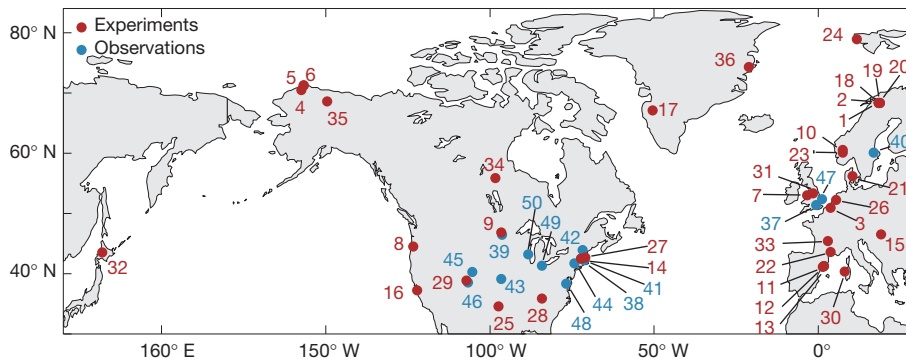
Predicting species' responses to climate change is a major challenge in ecology. Plants have been a focus of study because their responses can affect entire food webs, disturbance regimes and crucial ecosystem services, including pollination, carbon and nutrient cycling, and water supply<sup>6</sup>. Researchers have adopted two main approaches to predict plant responses to climate change. The first approach relies on observations over time or space to quantify sensitivities to temperature variability and change, including recent warming that is at least partly associated with greenhouse gas emissions<sup>3–5</sup>. The second approach relies on experiments that directly warm natural plant communities on a small scale<sup>2,7</sup>. Results from both methods suggest that advancing phenology (for example, earlier flowering and leafing) is one of the most sensitive plant responses to warming. Current estimates of changes in phenology are 1.9–3.3 days per °C for experiments<sup>1</sup> and 2.5–5 days per °C for observations<sup>3,8</sup>.

Warming experiments have been used to extrapolate to future climate conditions for more than 20 years<sup>7,9</sup>. This approach rests on the critical but little-tested assumption that plant responses to experimental warming match long-term responses to global warming. Testing this assumption is an important first step in assessing the utility of warming experiments for long-term forecasting and prediction. Recent studies have compared experimental and observational results at single sites<sup>2,10</sup>, and one study found coherence across methods in plant responses to warming<sup>10</sup>; however, cross-site comparisons have proved more challenging.

Here we present two new spatially and temporally extensive databases of plant phenology for 1,634 species based on long-term observations and short-term warming experiments (Fig. 1). The databases were developed specifically to test how accurately short-term warming experiments predict the long-term responses of wild plants. We tested for differences in the overall estimated phenological sensitivity to temperature (change in flowering or leafing date per °C, hereafter referred to as 'temperature sensitivity') between the experiments and the observations. The sensitivities for each species at a site,  $i$ , were calculated as the  $\beta$  coefficient from the following: phenological event date <sub>$i$</sub>  =  $\beta_i(\text{MAT}) + \varepsilon_i$ , where MAT is the mean annual temperature for each site and  $\varepsilon$  is error. Our observational sensitivities are thus based on interannual temperature variability. For the sensitivities from experiments, we used a regression approach similar to the observational calculation, which simplifies to the following: (phenological event date <sub>$i$ ,warm</sub> – phenological event date <sub>$i$ ,control</sub>) /  $\Delta T$ , where  $\Delta T$  is the change in the temperature reported between the warmed and the control plots (Supplementary Information). We looked at how the sensitivities to temperature varied across space and time, and we tested whether sensitivity increases with latitude<sup>5</sup> or with the magnitude of temperature change<sup>8</sup>. We also assessed whether species that flower relatively early in the spring are more sensitive to temperature than later-flowering species, as has been found in several observational<sup>3</sup> and experimental<sup>10,11</sup> studies. Although other abiotic factors, including photoperiod<sup>12</sup> and snowmelt<sup>13</sup>, may affect phenology, we focus on temperature here because it is the most consistent and dominant controller of spring phenology<sup>12,14</sup>, as well as the most reported.

Warming experiments underpredicted observational responses to climate change (Fig. 2). Across all species, the experiments underpredicted the magnitude of the advance—for both leafing and flowering—that results from temperature increases (Fig. 2a; flowering,  $F_{1,34} = 9.7$ ,

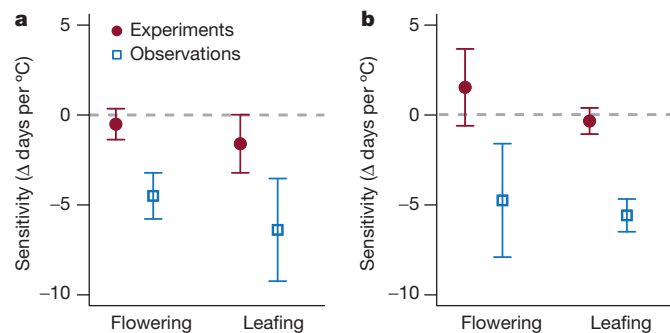
<sup>1</sup>Division of Biological Sciences, University of California San Diego, 9500 Gilman Drive 0116, La Jolla, California 92093, USA. <sup>2</sup>NASA Goddard Institute for Space Studies, New York, New York 10025, USA. <sup>3</sup>Ocean and Climate Physics, Lamont-Doherty Earth Observatory, Palisades, New York 10964-8000, USA. <sup>4</sup>Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269-3043, USA. <sup>5</sup>USA National Phenology Network, 1955 East Sixth Street, Tucson, Arizona 85721, USA. <sup>6</sup>US Geological Survey, 1955 East Sixth Street, Tucson, Arizona 85719, USA. <sup>7</sup>Department of Biological Sciences, North Dakota State University, Fargo, North Dakota 58108, USA. <sup>8</sup>National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, California 93101, USA. <sup>9</sup>Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, Quebec H3A 1B1, Canada. <sup>10</sup>Biodiversity Research Centre, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada. <sup>11</sup>Department of Biology, University of Maryland, College Park, Maryland 20742, USA. <sup>12</sup>National Center for Atmospheric Research, P.O. Box 3000, Boulder, Colorado 80307, USA. <sup>13</sup>Swedish University of Agricultural Sciences, Swedish National Phenology Network, Åsa, Unit for Field-based Forest Research, SE-36030 Lammhult, Sweden. <sup>14</sup>Theoretical Population Ecology and Evolution, Lund University, SE-22362 Lund, Sweden. <sup>15</sup>Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Santa Barbara, California 93106, USA. <sup>16</sup>US Geological Survey, Denver Federal Center, MS 412, Denver, Colorado 80225, USA. <sup>17</sup>School of Biology and Ecology & Sustainability Solutions Initiative, University of Maine, Orono, Maine 04469, USA. <sup>18</sup>Integrative Biology, University of Texas, 1 University Station C0930, Austin, Texas 78712, USA. <sup>19</sup>Marine Sciences Institute, A425 Portland Square, Drake Circus, University of Plymouth, Plymouth, Devon PL4 8AA, UK. <sup>20</sup>Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland. <sup>21</sup>Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland. <sup>22</sup>Department of Geography, Bolton 410, PO Box 413, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201-0413, USA.



**Figure 1 | Map of the experimental and observational sites.** The map covers 14 long-term observational phenology studies and 36 experimental phenology studies (one experimental study site, in Australia, is not shown). The numbers correspond to site information given in the Supplementary Information.

$P = 0.004$ ; leafing,  $F_{1,18} = 2.8$ ,  $P = 0.1$ ). The estimates for observations were based on plant responses to interannual temperature variability, yet they were similar to estimates calculated from long-term trends (Fig. 3) and were consistent across alternative approaches to calculating mean sensitivities (Supplementary Information). Because there are far more species represented in our observational data (1,558 unique taxa) than in the experimental data (115), this difference in mean temperature sensitivities could be attributable to a sampling effect, in which the increased sampling of species adds species that are more sensitive. However, when analyses were restricted to the 36 species common to both study types, the experiments still produced smaller sensitivities (in the sign tests for flowering and leafing,  $P = 0.02$ ,  $n = 30$  for flowering and  $n = 7$  for leafing). Most strikingly, the experiments predicted a delay in flowering and little change in leafing in response to higher temperatures, whereas the observations predicted an advance of  $>4.6$  days per  $^{\circ}\text{C}$  for either phenological event (Fig. 2b).

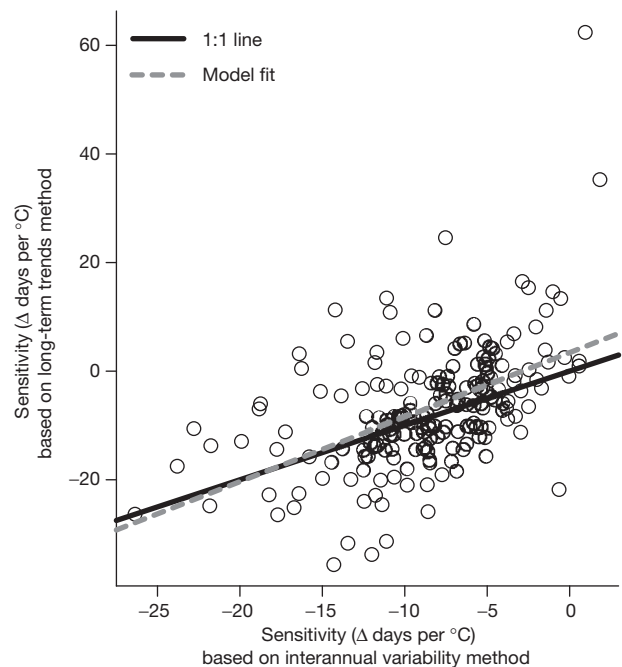
The experiments also failed to match the observational studies' predictions of greater advances in the flowering of early-season species with warming (Fig. 4; mean flowering date,  $F_{1,2114} = 93.2$ ,  $P < 0.0001$ ; flowering date  $\times$  study type,  $F_{1,2114} = 3.7$ ,  $P = 0.05$ ). Across experimental and observational study types, sensitivities were not associated with latitude ( $F_{1,31} = 0.02$ ,  $P > 0.8$ ) or interannual climate variability ( $F_{1,31} = 0.1$ ,  $P > 0.7$ ), suggesting that, although the interspecific variation in sensitivities is great, communities that are governed by different climate regimes do not differ in their overall sensitivity to warming. Additionally, we found no evidence that sensitivities varied with the degree of experimental warming ( $F_{1,76} = 1.1$ ,  $P = 0.3$ ); this is in contrast to the observational studies, which show an increase in sensitivities with



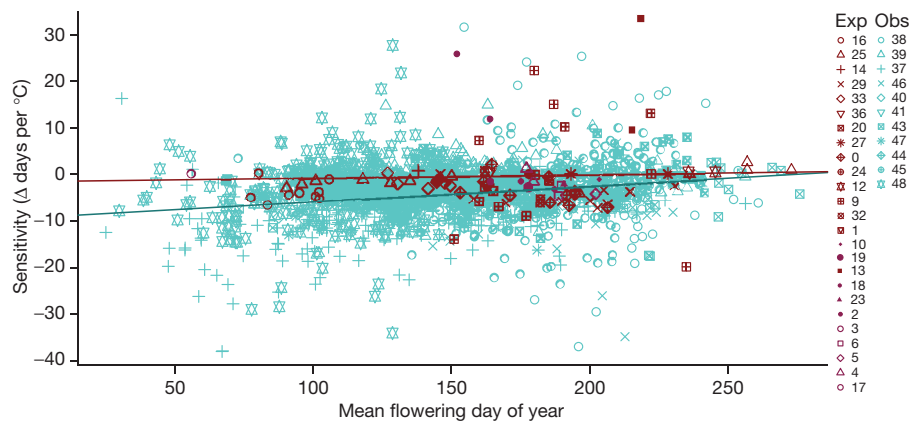
**Figure 2 | Estimates of the flowering and leafing sensitivities.** The estimates from the mixed effects model (presented as mean  $\pm$  s.e.m.), including the random effects of site and species, show that experiments underpredict the magnitude of plant responses to interannual temperature variation for all species sampled (a) and for the species that are common to both the experimental and the observational data sets (b). The region above the dashed grey line represents positive sensitivities, meaning that the species' phenological events are delayed with warming, whereas the region below the line represents negative sensitivities, meaning that the species' events advance with warming.

the degree of warming (Supplementary Information;  $F_{1,4856} = 12.3$ ,  $P = 0.0005$ ; see also ref. 8).

Several artefacts might explain the lower sensitivities in the experimental studies, including differences in the traits of the species sampled, correlations between MAT and other environmental variables for the observational studies, and differences in the degree of warming. We found no evidence, however, that these factors explain the smaller or positive sensitivities found in the experiments. Both study types mainly examined longer-lived plant species (85.0% and 88.4% were perennials in the experiments and the observations, respectively), and most were herbs or forbs (78.0% and 72.3%, respectively) not woody species (Supplementary Information). In addition, we found no correlations between MAT and other climate variables for which data were



**Figure 3 | Relationship between sensitivities to temperature calculated from interannual variation and from long-term trends.** Observational sensitivities calculated from interannual variation (regression of flowering date versus MAT) were strongly correlated with those calculated from long-term trends (long-term trend in the flowering date versus long-term trend in MAT). The mixed effects model fit ( $F_{1,185} = 63.9$ ,  $P < 0.0001$ ) and the 1:1 line are both shown. For long-term sensitivity estimates, we used the subset of the NECTAR data corresponding to species with observations between 1971 and 2000 (a common standard for long-term climate trend analysis). See the Supplementary Information (Observational sensitivities: comparisons with long-term trends and 1971–2000 standard) for complete details and for further alternative methods and data for calculating observational sensitivities.



**Figure 4 | Sensitivity to temperature decreases throughout the growing season.** This trend is seen in the observational studies (blue) but not in the experimental studies (red). The numbers correspond to those in Fig. 1 and to site information given in the Supplementary Information. The linear model

fitted to the observational data predicts that species blooming 3 months apart would have differences in sensitivities of  $>3$  days per  $^{\circ}\text{C}$ . Exp, experimental data; Obs, observational data.

available (Supplementary Information). Finally, both study types spanned roughly a  $0\text{--}5^{\circ}\text{C}$  temperature range, with a median of  $1.4\text{--}2.4^{\circ}\text{C}$  (Supplementary Information).

Given the difference in time frames between the experimental and observational studies (3.8 years and 31.0 years, respectively;  $F_{1,49} = 72.0$ ,  $P < 0.0001$ ), the experiments may capture predominantly plastic responses to temperature change, whereas the observational studies may also integrate shifts in genotypes, community composition and ecosystem dynamics. Furthermore, plant responses may have shifted fundamentally following the significant Northern Hemisphere warming in all seasons starting in the 1970s<sup>15</sup>, thus the experiments (which have been conducted relatively recently) may capture only these shifted dynamics. However, we found that these scaling issues would create a bias towards detecting greater sensitivities from short-term, post-1970 studies (Supplementary Information), which is opposite to the trends in the experimental studies. Sensitivities that captured 31–35 years of phenological change were generally smaller than those that spanned 5 or fewer years (sign tests of  $P \leq 0.02$ , Supplementary Information). In addition, the sensitivities were similar before 1970 and after 1970 ( $P > 0.1$ , Supplementary Information), when warming trends began to accelerate<sup>15</sup>.

Although warming methods have continually improved<sup>16–18</sup>, all experimental manipulations unavoidably alter additional environmental factors. For example, the most common passive warming structures—open-top chambers—reduce light, wind and often soil moisture and can unintentionally increase minimum winter temperatures<sup>19,20</sup>. By contrast, above-canopy heaters, which are common structures for active warming, often achieve warming only in periods of low wind<sup>17</sup>. We found no evidence of differences in the sensitivities between the passive and active designs ( $F_{1,23} = 0.05$ ,  $P > 0.8$ ). Designs using above-canopy heating, however, tended to produce the greatest phenological advances for flowering (Supplementary Information;  $F_{2,22} = 7.2$ ,  $P = 0.004$ ), and this was the type of heating used in the one study that found correspondence between experimental and observational methods<sup>10</sup> (Supplementary Information). There were no effects of habitat (for example, alpine or arctic), study duration or period of warming each year (Supplementary Information); the results from studies that manipulated temperature only during the growing season were indistinguishable from those that elevated temperature year round ( $F_{1,18} = 0.5$ ,  $P > 0.4$ ). Both active and passive experimental designs tend to reduce total irradiance and soil moisture<sup>11,19,21</sup>, which can delay plant phenology<sup>2,22</sup>. Thus, the lower sensitivities of the experiments could be the product of planned temperature increases that tend to advance phenology and artefacts that tend to delay phenology. Alternatively, the lower sensitivities of the experiments could occur if experimental studies

isolate temperature effects, whereas observational data integrate the effects of complex and possibly reinforcing interactions among multiple drivers of anthropogenic climate change and variability (Supplementary Information).

One needed improvement for both experimental and observational studies is the increased reporting of temperature and other relevant environmental metrics. Many experimental studies report coarse estimates of temperature differences (for example, annual means), often omitting measurement duration, frequency and statistics. Furthermore, the few efforts to examine temperature changes carefully have found that warming chambers can sometimes result in the cooling of plots<sup>20</sup> or can have unintended effects on daily and seasonal temperature regimes<sup>19</sup>. Such variation might confound experimental study data because plants are differentially sensitive to temperatures during different parts of the day and integrate temperature differentially over months and seasons<sup>23</sup>. Moving towards more relevant measures of temperature sensitivity should also be a future goal for observational studies. For example, using MAT, we found that 23.4% of the species in the observational data responded significantly to temperature, but this proportion nearly tripled (to 67.6%) when a growing-degree day model was used (Supplementary Information). Additionally, researchers collecting experimental or observational data should report other environmental metrics that may drive phenology. For example, spring phenology is driven by snowmelt date in many alpine and tundra habitats<sup>10,24</sup>. This date is not reported, however, at most sites, and experiments often remove warming structures in the snowpack season, thus not altering snowmelt date. Finally, we found high interspecific variation in sensitivities in both study types, suggesting that increased species sampling is needed for accurate climate change forecasting.

Our estimates of wild plant sensitivity to temperature, based on observations for 1,558 species, suggest that spring leafing and flowering will continue to advance at the rate of  $5\text{--}6$  days per  $^{\circ}\text{C}$  in many regions of the world (within the limits of historical temperature ranges)<sup>25</sup>. The database underlying our estimates represents a sampling of species that is almost an order of magnitude higher than the most recent comprehensive meta-analysis of plant phenology<sup>3</sup>; yet sensitivity estimates using data derived from ref. 3 are highly similar to those presented here (Supplementary Information). Comparisons with temporal estimates (for example, the change in days per decade<sup>5,26</sup>) are difficult; however, given a global increase in temperature of  $0.6^{\circ}\text{C}$  over the past four decades<sup>15</sup>, our estimates seem to be in line with the current estimates of  $-1.1$  to  $-3.3$  days per decade for plants<sup>5</sup>.

Accurate models of biotic responses to climate change require comparable methods to estimate how plants shift with increasing temperatures. Field-based warming studies allow us to mechanistically test single versus multiple factors and to project forward to novel global



conditions for which historical observations offer no comparisons<sup>27</sup>. To project experimental results, however, we must first verify that they match the changes that are observed in natural systems. Providing a first step, we have shown that warming experiments underpredict the advance of spring events observed over recent decades. Furthermore, when sampling the same species, the experiments failed to predict both the magnitude and the direction of plant responses to warming. Such differences between observed and experimentally estimated temperature sensitivities indicate that experimental results alone should not be used for parameterizing species distribution and ecosystem models. Although long-term observational data are currently the best measures of biotic responses to anthropogenic climate change<sup>28</sup>, our results suggest that researchers may not understand these responses well enough to replicate them experimentally. Obvious remedies include ongoing efforts to improve experimental designs and to expand observational networks, improved prediction of species-level variation, and refinement and standardization of temperature sensitivities across experimental and observational studies.

## METHODS SUMMARY

Complete methods are given in the Data and Methods sections of the Supplementary Information.

Received 21 November 2011; accepted 6 March 2012.

Published online 2 May 2012.

- Arft, A. M. *et al.* Responses of tundra plants to experimental warming: meta-analysis of the international tundra experiment. *Ecol. Monogr.* **69**, 491–511 (1999).
- Morin, X., Roy, J., Sonie, L. & Chuine, I. Changes in leaf phenology of three European oak species in response to experimental climate change. *New Phytol.* **186**, 900–910 (2010).
- Menzel, A. *et al.* European phenological response to climate change matches the warming pattern. *Glob. Change Biol.* **12**, 1969–1976 (2006).
- Schwartz, M. D., Ahas, R. & Aasa, A. Onset of spring starting earlier across the Northern Hemisphere. *Glob. Change Biol.* **12**, 343–351 (2006).
- Parmesan, C. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Glob. Change Biol.* **13**, 1860–1872 (2007).
- Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A. & Schwartz, M. D. Shifting plant phenology in response to global change. *Trends Ecol. Evol.* **22**, 357–365 (2007).
- Chapin, F. S. & Shaver, G. R. Physiological and growth responses of arctic plants to a field experiment simulating climatic change. *Ecology* **77**, 822–840 (1996).
- Amano, T., Smithers, R. J., Sparks, T. H. & Sutherland, W. J. A 250-year index of first flowering dates and its response to temperature changes. *Proc. R. Soc. Lond. B* **277**, 2451–2457 (2010).
- Harte, J. & Shaw, R. Shifting dominance within a montane vegetation community: results of a climate-warming experiment. *Science* **267**, 876–880 (1995).
- Dunne, J. A., Harte, J. & Taylor, K. J. Subalpine meadow flowering phenology responses to climate change: integrating experimental and gradient methods. *Ecol. Monogr.* **73**, 69–86 (2003).
- Sherry, R. A. *et al.* Divergence of reproductive phenology under climate warming. *Proc. Natl Acad. Sci. USA* **104**, 198–202 (2007).
- Chuine, I., Morin, X. & Bugmann, H. Warming, photoperiods, and tree phenology. *Science* **329**, 277–278 (2010).
- Dunne, J. A., Saleska, S. R., Fischer, M. L. & Harte, J. Integrating experimental and gradient methods in ecological climate change research. *Ecology* **85**, 904–916 (2004).
- Rohde, A., Bastien, C. & Boerjan, W. Temperature signals contribute to the timing of photoperiodic growth cessation and bud set in poplar. *Tree Physiol.* **31**, 472–482 (2011).
- Trenberth, K. E. & Josey, S. A. *Observations: Surface and Atmospheric Climate Change 235–336* (IPCC, 2007).
- Marion, G. M. *et al.* Open-top designs for manipulating field temperature in high-latitude ecosystems. *Glob. Change Biol.* **3**, 20–32 (1997).
- Kimball, B. A. Theory and performance of an infrared heater for ecosystem warming. *Glob. Change Biol.* **11**, 2041–2056 (2005).
- Kimball, B. A. *et al.* Infrared heater arrays for warming ecosystem field plots. *Glob. Change Biol.* **142**, 309–320 (2008).
- Kennedy, A. D. Simulated climate-change: are passive greenhouses a valid microcosm for testing the biological effects of environmental perturbations? *Glob. Change Biol.* **1**, 29–42 (1995).
- Molau, U. Responses to natural climatic variation and experimental warming in two tundra plant species with contrasting life forms: *Cassiope tetragona* and *Ranunculus nivalis*. *Glob. Change Biol.* **3**, 97–107 (1997).
- Kennedy, A. D. Temperature effects of passive greenhouse apparatus in high-latitude climate-change experiments. *Funct. Ecol.* **9**, 340–350 (1995).
- Rathcke, B. & Lacey, E. P. Phenological patterns of terrestrial plants. *Annu. Rev. Ecol. Syst.* **16**, 179–214 (1985).
- Aikawa, S., Kobayashi, M. J., Satake, A., Shimizu, K. K. & Kudoh, H. Robust control of the seasonal expression of the *Arabidopsis FLC* gene in a fluctuating environment. *Proc. Natl Acad. Sci. USA* **107**, 11632–11637 (2010).
- Chapin, F. S., Matson, P. A. & Mooney, H. A. *Principles of Terrestrial Ecosystem Ecology* (Springer, 2002).
- Körner, C. & Basler, D. Phenology under global warming. *Science* **327**, 1461–1462 (2010).
- Root, T. L. *et al.* Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57–60 (2003).
- Williams, J. W. & Jackson, S. T. Novel climates, no-analog communities, and ecological surprises. *Front. Ecol. Environ.* **59**, 475–482 (2007).
- Rosenzweig, C. *et al.* Attributing physical and biological impacts to anthropogenic climate change. *Nature* **453**, 353–357 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was conducted as part of the Forecasting Phenology Working Group supported by the National Center for Ecological Analysis & Synthesis (EF-0553768), with additional support from National Science Foundation grants DBI-0905806, IOS-0639794, DEB-0922080 and the Natural Sciences and Engineering Research Council of Canada CREATE Training Program. Special thanks to the many data managers, including G. Aldridge, P. Huth, D. Inouye, G. Johansson, A. Miller-Rushing, J. O'Keefe, R. Primack, S. Smiley, T. Sparks and J. Thompson. We thank M. Ayres, L. Kueppers, D. Moore and M. O'Connor for comments on earlier drafts.

**Author Contributions** E.M.W. conceived the idea, performed analyses and wrote the paper. B.I.C. performed analyses. E.E.C. and N.J.B.K. assisted with analyses. E.M.W., E.E.C., J.M.A., T.M.C., S.E.T. and S.P. developed the STONE database. E.M.W., B.I.C. and J.R. contributed extensively to development of the NECTAR database. All authors (including J.L.B., T.J.D., T.R.A., K.B., S.J.M., G.J.M., B.J.M., C.P., N.S. and M.D.S.) contributed to the editing of the manuscript and to data management of the observational data sets.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to E.M.W. ([wolkovich@biodiversity.ubc.ca](mailto:wolkovich@biodiversity.ubc.ca)).

# Prion-like behaviour and tau-dependent cytotoxicity of pyroglutamylated amyloid- $\beta$

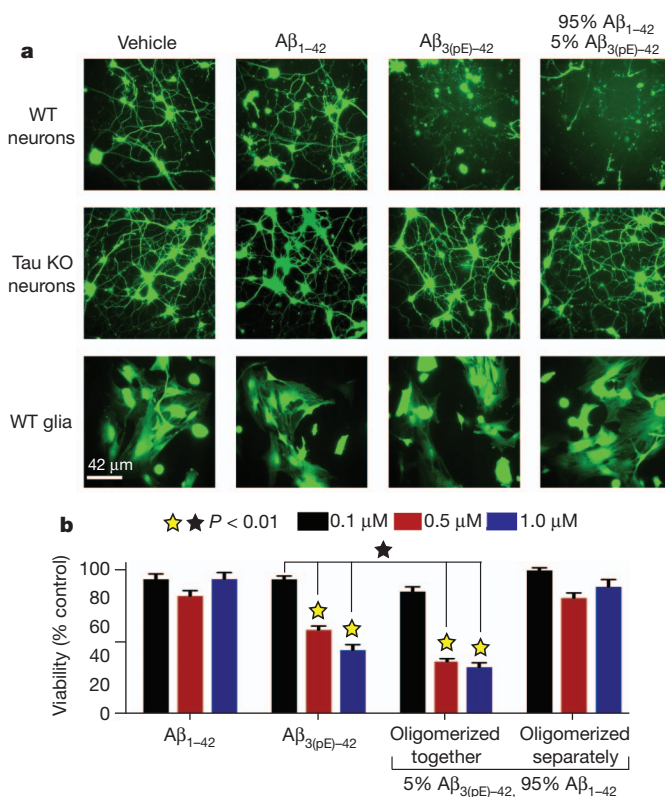
Justin M. Nussbaum<sup>1\*</sup>, Stephan Schilling<sup>2\*</sup>, Holger Cynis<sup>2</sup>, Antonia Silva<sup>1</sup>, Eric Swanson<sup>1</sup>, Tanaporn Wangsanut<sup>1</sup>, Kaycie Taylor<sup>3</sup>, Brian Wiltgen<sup>3</sup>, Asa Hatami<sup>4</sup>, Raik Röncke<sup>5</sup>, Klaus Reymann<sup>5</sup>, Birgit Hutter-Paier<sup>6</sup>, Anca Alexandru<sup>7</sup>, Wolfgang Jagla<sup>7</sup>, Sigrid Graubner<sup>7</sup>, Charles G. Glabe<sup>4</sup>, Hans-Ulrich Demuth<sup>2,7</sup> & George S. Bloom<sup>1,8</sup>

Extracellular plaques of amyloid- $\beta$  and intraneuronal neurofibrillary tangles made from tau are the histopathological signatures of Alzheimer's disease. Plaques comprise amyloid- $\beta$  fibrils that assemble from monomeric and oligomeric intermediates, and are prognostic indicators of Alzheimer's disease. Despite the importance of plaques to Alzheimer's disease, oligomers are considered to be the principal toxic forms of amyloid- $\beta$ <sup>1,2</sup>. Interestingly, many adverse responses to amyloid- $\beta$ , such as cytotoxicity<sup>3</sup>, microtubule loss<sup>4</sup>, impaired memory and learning<sup>5</sup>, and neuritic degeneration<sup>6</sup>, are greatly amplified by tau expression. Amino-terminally truncated, pyroglutamylated (pE) forms of amyloid- $\beta$ <sup>7,8</sup> are strongly associated with Alzheimer's disease, are more toxic than amyloid- $\beta$ , residues 1–42 ( $A\beta_{1-42}$ ) and  $A\beta_{1-40}$ , and have been proposed as initiators of Alzheimer's disease pathogenesis<sup>9,10</sup>. Here we report a mechanism by which pE- $A\beta$  may trigger Alzheimer's disease.  $A\beta_{3(pE)-42}$  co-oligomerizes with excess  $A\beta_{1-42}$  to form metastable low- $n$  oligomers (LNOs) that are structurally distinct and far more cytotoxic to cultured neurons than comparable LNOs made from  $A\beta_{1-42}$  alone. Tau is required for cytotoxicity, and LNOs comprising 5%  $A\beta_{3(pE)-42}$  plus 95%  $A\beta_{1-42}$  (5% pE- $A\beta$ ) seed new cytotoxic LNOs through multiple serial dilutions into  $A\beta_{1-42}$  monomers in the absence of additional  $A\beta_{3(pE)-42}$ . LNOs isolated from human Alzheimer's disease brain contained  $A\beta_{3(pE)-42}$ , and enhanced  $A\beta_{3(pE)-42}$  formation in mice triggered neuron loss and gliosis at 3 months, but not in a tau-null background. We conclude that  $A\beta_{3(pE)-42}$  confers tau-dependent neuronal death and causes template-induced misfolding of  $A\beta_{1-42}$  into structurally distinct LNOs that propagate by a prion-like mechanism. Our results raise the possibility that  $A\beta_{3(pE)-42}$  acts similarly at a primary step in Alzheimer's disease pathogenesis.

pE- $A\beta$  peptides contain an amino-terminal pyroglutamate, whose modification from glutamate is catalysed by glutaminyl cyclase (QC; also known as QPCT)<sup>10</sup>. The most prominent pE- $A\beta$  species *in vivo* are  $A\beta_{3(pE)-40}$ ,  $A\beta_{3(pE)-42}$ ,  $A\beta_{11(pE)-40}$  and  $A\beta_{11(pE)-42}$  (ref. 8; Supplementary Fig. 1), with  $A\beta_{3(pE)-42}$  being most abundant<sup>11</sup>. pE- $A\beta$  is more cytotoxic<sup>12</sup> and aggregates more rapidly<sup>13,14</sup> than conventional amyloid- $\beta$ , and QC activity and pE- $A\beta$  levels are increased several-fold in Alzheimer's disease brain<sup>10</sup>. Alzheimer's disease mouse models also indicate a role for pE- $A\beta$  in initiating pathology: oral administration of a QC inhibitor led to improved memory and learning, and reduced levels of pE- $A\beta$  and conventional amyloid- $\beta$ <sup>10</sup>. These data imply that pE- $A\beta$  potentiates the neurotoxicity of conventional amyloid- $\beta$ , but leave open the issue of molecular mechanisms. To address that issue, we compared oligomerization of  $A\beta_{3(pE)-42}$ ,  $A\beta_{1-42}$ , and mixtures of the peptides *in vitro*, and analysed responses of primary cultured neurons and glial cells (Supplementary Fig. 2) to the oligomers.

At 5  $\mu$ M peptide, 5% pE- $A\beta$  aggregated faster than  $A\beta_{3(pE)-42}$  or  $A\beta_{1-42}$  alone, based on thioflavin T fluorescence shifts<sup>15</sup> (Supplementary

Fig. 3). The ratio of optical densities at 450 nm versus 490 nm ( $OD_{450\text{ nm}}/OD_{490\text{ nm}}$ ) for  $A\beta_{3(pE)-42}$  rose and peaked more rapidly than for  $A\beta_{1-42}$ , but peaked at a  $\sim 25\%$  lower level. The fastest rise in the



**Figure 1 | Tau-dependent cytotoxicity of oligomers formed by co-incubation of  $A\beta_{3(pE)-42}$  and  $A\beta_{1-42}$ .** Primary mouse wild-type (WT) and tau-knockout (KO) forebrain neurons, and secondary cultures of wild-type mouse glia were treated for 12 h with  $A\beta_{1-42}$ ,  $A\beta_{3(pE)-42}$ , or 5%  $A\beta_{3(pE)-42}$  plus 95%  $A\beta_{1-42}$ , which were oligomerized for 24 h at 5  $\mu$ M before dilution into culture media. **a**, Cells were exposed to calcein-AM and imaged live by epifluorescence microscopy to assay viability<sup>16</sup>. Extensive death and detachment of cells were observed only for wild-type neurons treated with  $A\beta_{3(pE)-42}$  or the 5%  $A\beta_{3(pE)-42}$  plus 95%  $A\beta_{1-42}$ . **b**, Following peptide treatment, cell viability was analysed by the XTT plate reader assay<sup>17</sup>. Note the robust cytotoxicity of  $A\beta_{3(pE)-42}$  containing solutions at concentrations as low as 0.5  $\mu$ M, unless  $A\beta_{3(pE)-42}$  and  $A\beta_{1-42}$  were incubated separately during oligomerization ( $P < 0.01$ ; yellow stars signify statistical significance of the indicated bar graphs versus vehicle controls; black stars signify statistical significance between the indicated bar graph pairs; mean  $\pm$  standard error of the mean (s.e.m.),  $n = 9$  replicates from 3 independent experiments).

<sup>1</sup>Department of Biology, University of Virginia, Charlottesville, Virginia 22904, USA. <sup>2</sup>Probiobio AG, 06120 Halle (Saale), Germany. <sup>3</sup>Department of Psychology, University of Virginia, Charlottesville, Virginia 22904, USA. <sup>4</sup>Department of Biochemistry and Molecular Biology, University of California at Irvine, Irvine, California 92697, USA. <sup>5</sup>Deutsches Zentrum fuer Neurodegenerative Erkrankungen, c/o Leibniz-Institut fuer Neurobiologie, 39118 Magdeburg, Germany. <sup>6</sup>JSW Life Sciences GmbH, A-8074 Grambach, Austria. <sup>7</sup>Ingenium Pharmaceuticals GmbH, 82152 Munich/Martinsried, Germany. <sup>8</sup>Department of Cell Biology, University of Virginia, Charlottesville, Virginia 22904, USA.

\*These authors contributed equally to this work.

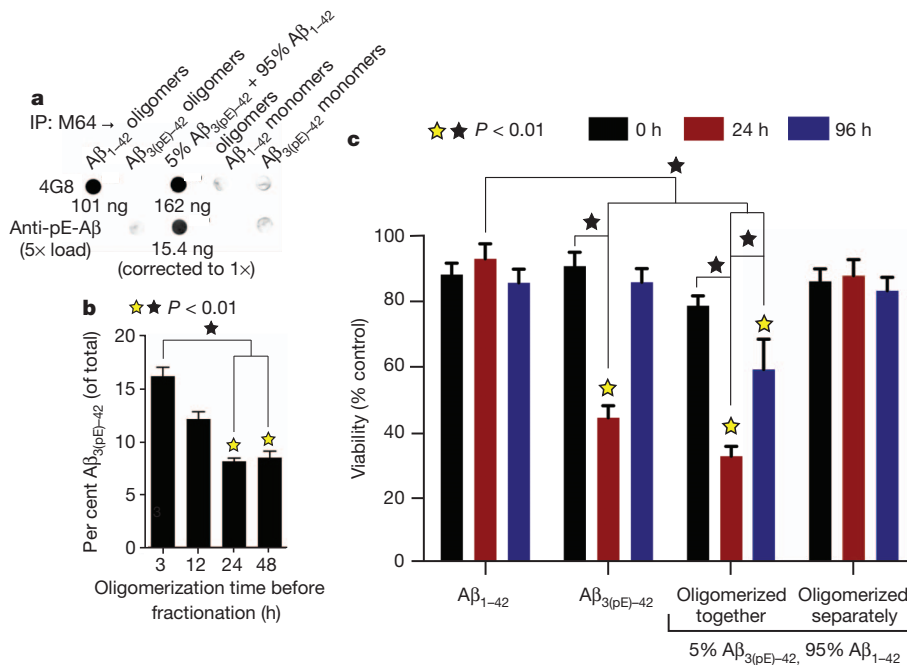
OD<sub>450 nm</sub>/OD<sub>490 nm</sub> ratio was for 5% pE-A $\beta$ , which peaked similarly to A $\beta$ <sub>3(pE)-42</sub>, A $\beta$ <sub>3(pE)-42</sub>, A $\beta$ <sub>1-42</sub> and 5% pE-A $\beta$  thus oligomerized by different pathways.

To test whether distinct biological activities were coupled to these oligomerization differences, we compared cytotoxicity of the peptides towards cultured neurons or glia using calcein-AM and fluorescence microscopy<sup>16</sup>. Twelve hours of A $\beta$ <sub>1-42</sub> exposure had little effect on cell viability for wild-type or tau-knockout neurons, or wild-type glial cells (Fig. 1a). Contrastingly, most wild-type neurons died and detached from the substrate after exposure to A $\beta$ <sub>3(pE)-42</sub> or 5% pE-A $\beta$ . Tau-knockout neurons and wild-type glia, which express little tau, were resistant to A $\beta$ <sub>3(pE)-42</sub> and 5% pE-A $\beta$ .

Cytotoxicity dose dependence was examined by incubating wild-type neurons for 24 h in oligomers comprising 0.1, 0.5 or 1  $\mu$ M peptides, and using the 2,3-bis-(2-methoxy-4-nitro-5-sulphophenyl)-2H-tetrazolium-5-carboxanilide (XTT) reduction assay<sup>17</sup> (Fig. 1b). Cells were unaffected by A $\beta$ <sub>1-42</sub>, but A $\beta$ <sub>3(pE)-42</sub> and 5% pE-A $\beta$  had substantial cytotoxicity at 0.5  $\mu$ M and even more at 1.0  $\mu$ M. Cytotoxicity of 5% pE-A $\beta$  required A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> to incubate together for 24 h before being added to cells. When they were incubated separately for 24 h and mixed together at a 1:19 molar ratio immediately before being applied to cells, they were not cytotoxic. A small amount of A $\beta$ <sub>3(pE)-42</sub> can thus markedly enhance the cytotoxicity of a large excess of A $\beta$ <sub>1-42</sub>, provided the two peptides oligomerize together.

Evidence for hybrid oligomers came from immunoprecipitation of various forms of amyloid- $\beta$  using aggregation-dependent M64, which does not recognize A $\beta$ <sub>3(pE)-42</sub> (see Supplementary Fig. 4 for characterization of all anti-amyloid- $\beta$  antibodies used, including M64). Immunoprecipitations were analysed on dot blots using 4G8, which equally recognizes A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub>, and anti-pE-A $\beta$ , which does not react with A $\beta$ <sub>1-42</sub>. M64 immunoprecipitated oligomers made from A $\beta$ <sub>1-42</sub> or 5% pE-A $\beta$ , but it did not immunoprecipitate A $\beta$ <sub>3(pE)-42</sub> oligomers, nor monomers of either peptide (Fig. 2a). Because anti-pE-A $\beta$  reacted with material immunoprecipitated out of 5% pE-A $\beta$ , M64 pulled down hybrid peptide oligomers. A $\beta$ <sub>3(pE)-42</sub> accounted for ~16% of the amyloid- $\beta$  in gel-filtered cytotoxic oligomers after 3 h of oligomerization, and steadily dropped to ~8% by 24 h (Fig. 2b). A $\beta$ <sub>3(pE)-42</sub> thus acts as a template that initiates formation of cytotoxic oligomers.

Cytotoxicity was sensitive to oligomerization time (Fig. 2c). Baseline cytotoxicity was observed at all time points for A $\beta$ <sub>1-42</sub>, and for 5% pE-A $\beta$  solutions in which A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> oligomerized separately. Pure A $\beta$ <sub>3(pE)-42</sub> killed ~50% of the cells after 24 h of oligomerization, but was virtually non-toxic at 0 h and after 96 h of oligomerization. The most cytotoxic solutions were 5% pE-A $\beta$ , in which the constituent peptides co-oligomerized for 24 h. These solutions killed ~60% of the cells within 24 h, and lower but robust cytotoxicity was observed at 96 h. Even the 0 h co-oligomers of 5% pE-A $\beta$  exhibited low, significant cytotoxicity. Co-incubated mixtures of 5% A $\beta$ <sub>3(pE)-42</sub> and 95% A $\beta$ <sub>1-42</sub>



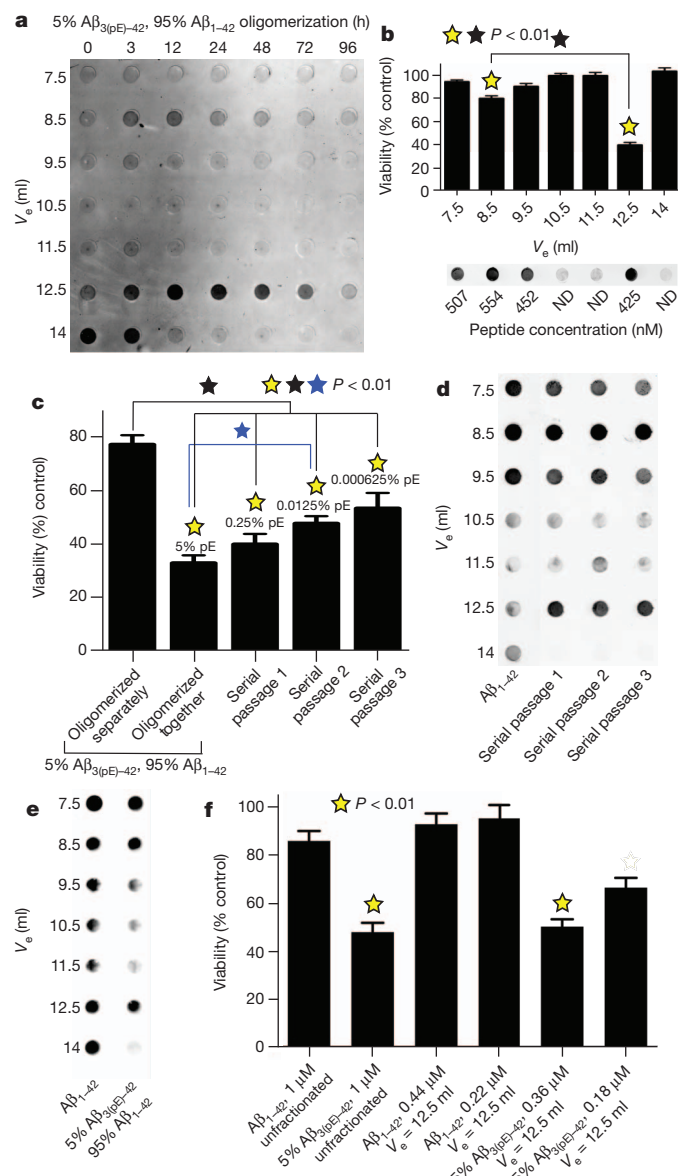
**Figure 2 | A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> form metastable, cytotoxic, hybrid oligomers.** **a**, A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> were incubated together at a 1:19 molar ratio (5% pE-A $\beta$ ) for 24 h at 1  $\mu$ M total amyloid- $\beta$ , and were then immunoprecipitated (IP) with M64, a rabbit monoclonal antibody that specifically recognizes residues 3–7 (EFRH) of A $\beta$ <sub>1-40</sub> oligomers or fibrils. Additional samples that were immunoprecipitated included otherwise identically treated oligomers made from pure A $\beta$ <sub>3(pE)-42</sub> or A $\beta$ <sub>1-42</sub>, and monomeric versions of the two peptides. Immunoprecipitated oligomers were converted to monomers by lyophilization, solubilization with HFIP and dilution into PBS, and along with the other samples were dot blotted onto nitrocellulose and analysed using 4G8, a mouse monoclonal antibody that recognizes A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> equally well, and an antibody that specifically recognizes pE-A $\beta$  (see Supplementary Fig. 4 for characterization of all antibodies used here). Quantification of the dot blots using a LI-COR Odyssey imaging station indicated that the oligomers that were immunoprecipitated from the mixed peptide solution contained both A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub>, at a molar ratio of ~1:10. **b**, Solutions containing 5% pE-A $\beta$ <sub>3(pE)-42</sub> and 95% A $\beta$ <sub>1-42</sub> were incubated for the indicated times, and then were fractionated by gel

filtration. At each time point, fractions that eluted at 12.5 ml, where most cytotoxicity resided (see Fig. 3b) were immunoprecipitated using anti-human amyloid- $\beta$  (N), an amino-terminal-specific antibody that does not react with pE-A $\beta$  (data not shown). The immunoprecipitates were then lyophilized, resolubilized with HFIP, and quantitatively analysed on dot blots with 4G8 and anti-pE-A $\beta$  using the LI-COR Odyssey. The time-dependent decrease in the A $\beta$ <sub>3(pE)-42</sub> content of the immunoprecipitated oligomers implies that A $\beta$ <sub>3(pE)-42</sub> initiated formation of hybrid peptide oligomers. **c**, A $\beta$ <sub>3(pE)-42</sub> and A $\beta$ <sub>1-42</sub> oligomerized for 0, 24 and 96 h either separately or together as 1:19 mixtures, and then were added to primary wild-type neuron cultures for 24 h at a final concentration of 1  $\mu$ M total amyloid- $\beta$ . Following peptide treatment, cell viability was analysed by the XTT plate reader assay<sup>17</sup>. The most cytotoxic species observed were the hybrid oligomers after 24 h of oligomerization ( $P < 0.01$ ; yellow stars signify statistical significance of the indicated bar graphs versus vehicle controls; black stars signify statistical significance between the indicated bar graph pairs; mean  $\pm$  s.e.m.,  $n = 6$  or 9 replicates from 3 independent experiments for panel **b** or **c**, respectively).



$A\beta_{1-42}$  can therefore form oligomers whose cytotoxicity is both greater and more enduring than oligomers formed by  $A\beta_{3(pE)-42}$  alone.

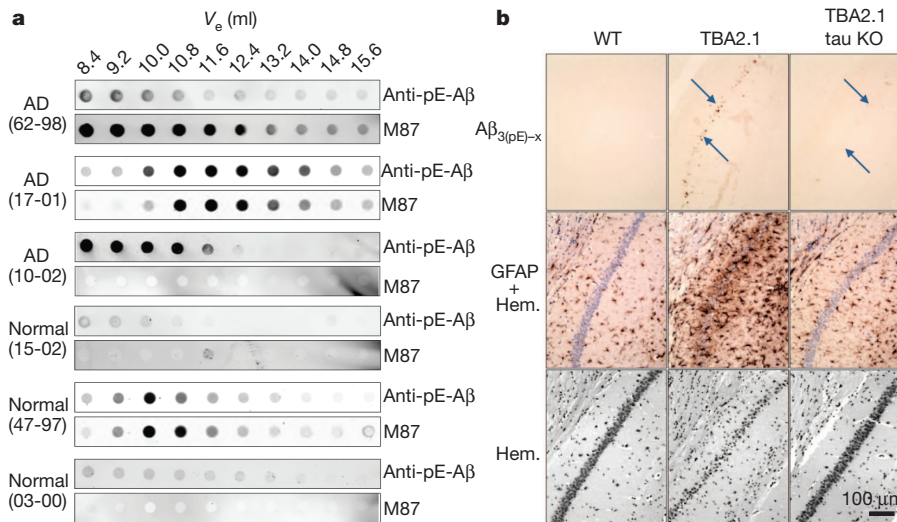
To identify the co-oligomer size(s) that were cytotoxic, amyloid- $\beta$  solutions were oligomerized for various times from 0–96 h before fractionation by gel filtration. Total amyloid- $\beta$  in all fractions was determined using 4G8 dot blots that, as shown in Fig. 3a (for 5% pE- $A\beta$ ) and Supplementary Fig. 5 (for  $A\beta_{1-42}$  and  $A\beta_{3(pE)-42}$ ), illustrate the full fractionation range of the column but exclude most void volume fractions. Presumptive monomeric  $A\beta_{1-42}$  dominated initially and persisted at 3 h, but was nearly undetectable after 12 h. The 3 h time point also marked the appearance of  $A\beta_{1-42}$  oligomers, which gradually increased in size over the next 93 h.  $A\beta_{3(pE)-42}$  and 5% pE- $A\beta$  oligomerized differently. Putative monomers were present at 0 h for both samples, when slightly larger species, LNOs that possibly corresponded to dimers or trimers (Supplementary Fig. 6), were also present. These persisted as the main species for 24 h for  $A\beta_{3(pE)-42}$  and for nearly 72 h for 5% pE- $A\beta$ , and later time points were dominated by larger aggregates that eluted in void volume fractions. Cytotoxicity was assayed for individual fractions of 5% pE- $A\beta$  that oligomerized for 24 h (Fig. 3b). Most cytotoxicity was associated with the possible dimers/trimers that eluted at 12.5 ml, which at 425 nM peptide killed more than 60% of the cells. Low cytotoxicity was also observed at 554 nM peptide for the larger oligomers that eluted at 8.5 ml.



The marked enhancement of  $A\beta_{1-42}$  cytotoxicity by  $A\beta_{3(pE)-42}$  suggested a prion-like templating mechanism of  $A\beta_{1-42}$  misfolding initiated by  $A\beta_{3(pE)-42}$ . To test that hypothesis, 5% pE- $A\beta$  that oligomerized for 24 h was diluted into 19 volumes of monomeric  $A\beta_{1-42}$ . A 24 h incubation of this mixture yielded 'serial passage 1', which was followed by two equivalent, sequential dilutions into monomeric  $A\beta_{1-42}$  to yield serial passages 2 and 3. A gradual loss of cytotoxicity was observed with successive passages, but even passage 3, which contained only 0.000625%  $A\beta_{3(pE)-42}$ , killed ~50% of the neurons within 24 h (Fig. 3c). Serially passaged gel-filtration samples contained abundant material that eluted at 12.5 ml in passages 1–3, despite the progressive dilution of  $A\beta_{3(pE)-42}$  (Fig. 3d).  $A\beta_{3(pE)-42}$  can therefore template formation of metastable, cytotoxic LNOs from excess  $A\beta_{1-42}$ , yielding potent bioactivity that can be serially passaged multiple times into monomeric  $A\beta_{1-42}$  without further addition of  $A\beta_{3(pE)-42}$ .

One possible explanation for why  $A\beta_{1-42}$  LNOs were inert is that they lacked sufficient properly sized oligomers. Accordingly, we altered the oligomerization protocol from 5  $\mu$ M peptide for 24 h at 37 °C to 10  $\mu$ M peptide for 30 min at 4 °C to obtain abundant  $A\beta_{1-42}$  oligomers that eluted at 12.5 ml (Fig. 3e). These LNOs were not cytotoxic (Fig. 3f), implying that they were structurally distinct from the putative dimers/trimers initiated by  $A\beta_{3(pE)-42}$ . This was confirmed by dot blots using M87, a conformation-sensitive anti-amyloid- $\beta$  antibody, to compare the putative dimers/trimers used for the cytotoxicity assays shown in Fig. 3f. We first lyophilized aliquots of all the amyloid- $\beta$  solutions, resuspended them with hexafluoroisopropanol (HFIP) to restore them to monomers, and then analysed them using 4G8. When parallel samples that were not lyophilized but were otherwise identical were analysed using M87, immunoreactivity was approximately twice as strong with LNOs made from  $A\beta_{1-42}$  versus those made from 5% pE- $A\beta$  (Supplementary Fig. 7). Cytotoxic LNOs of 5% pE- $A\beta$  are thus structurally distinct from comparably sized LNOs of  $A\beta_{1-42}$ .

**Figure 3 | The cytotoxic species are low-n, prion-like oligomers.** **a**, Gel-filtration chromatography was used to fractionate 5% pE- $A\beta$  after oligomerization at 5  $\mu$ M at 37 °C for 0–96 h. The resulting fractions were then converted to monomers using HFIP and analysed on dot blots using monoclonal antibody 4G8. Note the metastable oligomers with an average elution volume ( $V_e$ ) of 12.5 ml. **b**, Isolated gel-filtration fractions from the 24 h time point were added to wild-type neuron cultures for 24 h, after which the cells were assayed for cell viability using XTT<sup>17</sup>. Robust cytotoxicity was associated only with the  $V_e$  = 12.5 ml fraction, although the  $V_e$  = 8.5 ml fraction had low, but statistically significant cell killing activity ( $P < 0.01$ ; mean  $\pm$  s.e.m.,  $n$  = 9 replicates from 3 independent experiments). ND, not detected. **c**, Cytotoxic hybrid oligomers made by co-incubating a 1:19 ratio of  $A\beta_{3(pE)-42}$ : $A\beta_{1-42}$  for 24 h at 5  $\mu$ M were diluted into a 19-fold molar excess of freshly dissolved, monomeric  $A\beta_{1-42}$ , which was then incubated at 5  $\mu$ M for another 24 h to yield serial passage 1. Two further iterations of this strategy yielded serial passages 2 and 3. The starting material and its serially passaged derivatives were added to wild-type neurons at 1  $\mu$ M peptide for 24 h, after which cells were analysed using the XTT assay for cell viability<sup>17</sup>. Only a gradual loss of cytotoxicity was observed with each successive serial passage. **d**, Each serially passaged sample, as well as otherwise identically prepared oligomers made from pure  $A\beta_{1-42}$ , were fractionated by gel filtration and analysed on dot blots with 4G8. Note that all serially passaged samples contained metastable LNOs of  $V_e$  = 12.5 ml, which were absent from the pure  $A\beta_{1-42}$  samples. **e**,  $A\beta_{1-42}$  (10  $\mu$ M) that was oligomerized for 30 min at 4 °C, and 5%  $A\beta_{3(pE)-42}$  plus 95%  $A\beta_{1-42}$  (5  $\mu$ M) that was oligomerized for 24 h at 37 °C were fractionated by gel filtration and analysed on dot blots exactly using 4G8. Note the isolation of fractions with  $V_e$  = 12.5 ml from both preparations. **f**, Wild-type neurons were assayed for viability using the XTT plate reader assay<sup>17</sup> following 24 h of exposure to the indicated amyloid- $\beta$  preparations. Note the minimal cytotoxicity of unfractionated  $A\beta_{1-42}$  and  $A\beta_{1-42}$  with  $V_e$  = 12.5 ml ( $P < 0.01$ , mean  $\pm$  s.e.m.,  $n$  = 9 replicates from 3 independent experiments).  $P < 0.01$ ; yellow stars signify statistical significance of the indicated bar graphs versus vehicle controls; black stars and blue stars signify statistical significance between the indicated bar graph pairs; blue mean  $\pm$  s.e.m.,  $n$  = 9 replicates from 3 independent experiments for panels **b**, **c** and **f**.



**Figure 4** |  $A\beta_{3(pE)-42}$  *in vivo*. **a**, Cytosol obtained from human Alzheimer's disease (AD) and similarly aged normal brains (Supplementary Fig. 9) were fractionated by gel filtration, and analysed by dot blotting with anti-pE-A $\beta$  and M87. Note the appearance of pE-A $\beta$  in LNO fractions, including those that eluted at 12.4 ml, especially in the Alzheimer's disease samples. **b**, Three-

Several lines of evidence demonstrate *in vivo* relevance for the data described so far. First, we identified LNOs containing  $A\beta_{3(pE)-42}$  in three out of three Alzheimer's disease samples, based on gel filtration of human brain extracts followed by dot blots of resulting fractions with anti-pE-A $\beta$  and M87. In contrast, only one of three age-matched samples with normal neuropathological diagnoses was positive for  $A\beta_{3(pE)-42}$  (Fig. 4a and Supplementary Fig. 8). Second, we crossed TBA2.1 mice<sup>18</sup> into a tau-knockout background<sup>19</sup>. By 3 months, TBA2.1 mice accumulated small amounts (40–100 ng g<sup>-1</sup> brain weight) of  $A\beta_{3(pE)-42}$ , which formed primarily intraneuronal aggregates, and was associated with massive hippocampal neuron loss and gliosis<sup>18</sup>. Knocking out tau provided almost complete protection against neuron loss and glial activation (Fig. 4b). Additional *in vivo* data are shown in Supplementary Fig. 9. Long-term potentiation (LTP) of mouse hippocampal neurons in slice cultures was potently and equally inhibited by oligomers made from 5%  $A\beta_{3(pE)-42}$  or 100%  $A\beta_{3(pE)-42}$ , whereas  $A\beta_{1-42}$  oligomers had no effect on LTP. 1%  $A\beta_{3(pE)-42}$  provoked mild, but statistically insignificant LTP impairment (Supplementary Fig. 9a). To evaluate the effects of increased  $A\beta_{3(pE)-42}$  in animal models, we crossed mice with neuron-specific expression of human  $\beta$ -amyloid precursor protein (APP) harbouring Swedish and London mutations (hAPP<sub>SL</sub>)<sup>20</sup>, with mice expressing human QC<sup>21</sup>. Nine-month-old double (hAPP<sub>SL</sub>/hQC) and single (hAPP<sub>SL</sub>) transgenic mice were indistinguishable in terms of insoluble and soluble  $A\beta_{x-42}$  levels, but the double transgenics had approximately twofold more insoluble  $A\beta_{3(pE)-42}$  and approximately ninefold more soluble  $A\beta_{3(pE)-42}$  than single transgenics (Supplementary Fig. 9b). Further analysis of the soluble  $A\beta_{x-42}$  by the A4 assay<sup>22</sup> revealed an approximately eightfold excess of oligomers in the double versus single transgenics (Supplementary Fig. 9c). Double transgenics performed more poorly in Morris water maze tests (Supplementary Fig. 9d) and had reduced hippocampal immunoreactivity for the synapse marker, synaptophysin (Supplementary Fig. 9e). Finally, peri-hippocampal injection of 5% pE-A $\beta$  at 5  $\mu$ M into APP<sub>SwDI</sub>/NOS2<sup>-/-</sup> Alzheimer's disease model mice<sup>23</sup> led 3–5 months later to the presence of plaques containing both pE-A $\beta$  and conventional amyloid- $\beta$ . Comparable plaques were rarely seen in sham-injected Alzheimer's disease mice or in wild-type mice injected with 5% pE-A $\beta$  (Supplementary Fig. 9f). These collective *in vivo* results emphasize the physiological significance of the companion biochemical and cultured cell results.

month-old TBA2.1 mice generating  $A\beta_{3(pE)-42}$ <sup>18</sup> show amyloid- $\beta$  deposits (arrows), massive astrogliosis (GFAP) and neuron loss (Hem., haematoxylin nuclear staining), none of which are evident in comparably aged wild-type (WT) mice or TBA2.1/tau-knockout (KO) hybrids.

Our studies provide new insights into Alzheimer's disease pathogenesis by demonstrating that hypertoxic amyloid- $\beta$  oligomers can be triggered by small quantities of a specifically truncated and post-translationally modified version of amyloid- $\beta$ . Although some previous studies demonstrated that pE modification of amyloid- $\beta$  considerably enhances its aggregation kinetics<sup>13,14,24</sup>, toxicity<sup>12,18,25</sup> and resistance to degradation<sup>12</sup>, a mechanistic explanation for the unique properties of pE-A $\beta$  has been lacking until now. Prior studies suggest coincident appearance of  $A\beta_{3(pE)-42}$  with development or progression of human Alzheimer's disease<sup>26,27</sup>. Co-localization of QC and  $A\beta_{3(pE)-42}$  was found in cored plaques of vulnerable regions in Alzheimer's disease, and evidence was provided for axonal transport of  $A\beta_{3(pE)-x}$  from QC-rich neuronal populations of the entorhinal cortex and locus coeruleus<sup>28</sup>. As LNOs containing  $A\beta_{3(pE)-42}$  are reasonably stable (Fig. 3a), they might initiate tau-dependent cytotoxicity intracellularly during axonal transport<sup>29</sup> or extracellularly following release at remote hippocampal synapses<sup>30</sup> of projection neurons<sup>28</sup>. The  $A\beta_{3(pE)-42}$ -induced formation of toxic mixed oligomers provides a rationale for these previous observations, and the tau-dependent cytotoxicity of 5% pE-A $\beta$  establishes a new functional connection between amyloid- $\beta$  and tau in Alzheimer's disease pathogenesis.

## METHODS SUMMARY

Full descriptions of thioflavin T assays, cell culture, cell viability assays, procedures for oligomerization of amyloid- $\beta$  peptides and their fractionation by gel-filtration chromatography, production and specificity of rabbit monoclonal anti-amyloid- $\beta$  antibodies, immunoprecipitation, dot blots and western blots, generation of hAPP<sub>SL</sub>/hQC transgenic mice, LTP measurements of mouse hippocampal slice cultures, peri-hippocampal injection of 5% pE-A $\beta$  into Alzheimer's disease model mice, cultured cell and brain immunohistochemistry, and collection of human brain extracts are provided in Supplementary Methods.

Received 16 May 2011; accepted 16 March 2012.

Published online 2 May 2012.

- Gandy, S. *et al.* Days to criterion as an indicator of toxicity associated with human Alzheimer amyloid- $\beta$  oligomers. *Ann. Neurol.* **68**, 220–230 (2010).
- Walsh, D. M. & Selkoe, D. J.  $A\beta$  oligomers—a decade of discovery. *J. Neurochem.* **101**, 1172–1184 (2007).
- Rapoport, M., Dawson, H. N., Binder, L. I., Vitek, M. P. & Ferreira, A. Tau is essential to  $\beta$ -amyloid-induced neurotoxicity. *Proc. Natl Acad. Sci. USA* **99**, 6364–6369 (2002).
- King, M. E. *et al.* Tau-dependent microtubule disassembly initiated by pre-fibrillar  $\beta$ -amyloid. *J. Cell Biol.* **175**, 541–546 (2006).

5. Roberson, E. D. *et al.* Reducing endogenous tau ameliorates amyloid  $\beta$ -induced deficits in an Alzheimer's disease mouse model. *Science* **316**, 750–754 (2007).
6. Jin, M. *et al.* Soluble amyloid  $\beta$ -protein dimers isolated from Alzheimer cortex directly induce Tau hyperphosphorylation and neuritic degeneration. *Proc. Natl Acad. Sci. USA* **108**, 5819–5824 (2011).
7. Mori, H., Takio, K., Ogawara, M. & Selkoe, D. J. Mass spectrometry of purified amyloid  $\beta$  protein in Alzheimer's disease. *J. Biol. Chem.* **267**, 17082–17086 (1992).
8. Saido, T. C. *et al.* Dominant and differential deposition of distinct  $\beta$ -amyloid peptide species,  $A\beta_{N3(pE)}$ , in senile plaques. *Neuron* **14**, 457–466 (1995).
9. Jawhar, S., Wirths, O. & Bayer, T. A. Pyroglutamate amyloid- $\beta$  (A $\beta$ ): a hatchet man in Alzheimer disease. *J. Biol. Chem.* **286**, 38825–38832 (2011).
10. Schilling, S. *et al.* Glutaminyl cyclase inhibition attenuates pyroglutamate A $\beta$  and Alzheimer's disease-like pathology. *Nature Med.* **14**, 1106–1111 (2008).
11. Tabaton, M. *et al.* Soluble amyloid  $\beta$ -protein is a marker of Alzheimer amyloid in brain but not in cerebrospinal fluid. *Biochem. Biophys. Res. Commun.* **200**, 1598–1603 (1994).
12. Russo, C. *et al.* Pyroglutamate-modified amyloid  $\beta$ -peptides—A $\beta$ N3(pE)—strongly affect cultured neuron and astrocyte survival. *J. Neurochem.* **82**, 1480–1489 (2002).
13. Schilling, S. *et al.* On the seeding and oligomerization of pGlu-amyloid peptides (*in vitro*). *Biochemistry* **45**, 12393–12399 (2006).
14. Schlenzig, D. *et al.* Pyroglutamate formation influences solubility and amyloidogenicity of amyloid peptides. *Biochemistry* **48**, 7072–7078 (2009).
15. Levine, H. III. Thioflavine T interaction with synthetic Alzheimer's disease  $\beta$ -amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci.* **2**, 404–410 (1993).
16. Wang, X. M. *et al.* A new microcellular cytotoxicity test based on calcein AM release. *Hum. Immunol.* **37**, 264–270 (1993).
17. Scudiero, D. A. *et al.* Evaluation of a soluble tetrazolium/formazan assay for cell growth and drug sensitivity in culture using human and other tumor cell lines. *Cancer Res.* **48**, 4827–4833 (1988).
18. Alexandru, A. *et al.* Selective hippocampal neurodegeneration in transgenic mice expressing small amounts of truncated A $\beta$  is induced by pyroglutamate-A $\beta$  formation. *J. Neurosci.* **31**, 12790–12801 (2011).
19. Tucker, K. L., Meyer, M. & Barde, Y. A. Neurotrophins are required for nerve growth during development. *Nature Neurosci.* **4**, 29–37 (2001).
20. Rockenstein, E., Mallory, M., Mante, M., Sisk, A. & Masliah, E. Early formation of mature amyloid- $\beta$  protein deposits in a mutant APP transgenic model depends on levels of A $\beta_{1-42}$ . *J. Neurosci. Res.* **66**, 573–582 (2001).
21. Jawhar, S. *et al.* Overexpression of glutaminyl cyclase, the enzyme responsible for pyroglutamate A $\beta$  formation, induces behavioral deficits, and glutaminyl cyclase knock-out rescues the behavioral phenotype in 5XFAD mice. *J. Biol. Chem.* **286**, 4454–4460 (2011).
22. Tanghe, A. *et al.* Pathological hallmarks, clinical parallels, and value for drug testing in Alzheimer's disease of the APP[V717I] London transgenic mouse model. *Int. J. Alzheimers Dis.* **2010** (2010).
23. Wilcock, D. M. *et al.* Progression of amyloid pathology to Alzheimer's disease pathology in an amyloid precursor protein transgenic mouse model by removal of nitric oxide synthase 2. *J. Neurosci.* **28**, 1537–1545 (2008).
24. He, W. & Barrow, C. J. The A $\beta$  3-pyroglutamyl and 11-pyroglutamyl peptides found in senile plaque have greater  $\beta$ -sheet forming and aggregation propensities *in vitro* than full-length A $\beta$ . *Biochemistry* **38**, 10871–10877 (1999).
25. Wirths, O. *et al.* Intraneuronal pyroglutamate-A $\beta$  3–42 triggers neurodegeneration and lethal neurological deficits in a transgenic mouse model. *Acta Neuropathol.* **118**, 487–496 (2009).
26. G ntert, A., Dobeli, H. & Bohrmann, B. High sensitivity analysis of amyloid- $\beta$  peptide composition in amyloid deposits from human and PS2APP mouse brain. *Neuroscience* **143**, 461–475 (2006).
27. Piccini, A. *et al.*  $\beta$ -Amyloid is different in normal aging and in Alzheimer disease. *J. Biol. Chem.* **280**, 34186–34192 (2005).
28. Hartlage-R bsamen, M. *et al.* Glutaminyl cyclase contributes to the formation of focal and diffuse pyroglutamate (pGlu)-A $\beta$  deposits in hippocampus via distinct cellular mechanisms. *Acta Neuropathol.* **121**, 705–719 (2011).
29. Vossel, K. A. *et al.* Tau reduction prevents A $\beta$ -induced defects in axonal transport. *Science* **330**, 198 (2010).
30. Wilcox, K. C., Lacor, P. N., Pitt, J. & Klein, W. L. A $\beta$  oligomer-induced synapse degeneration in Alzheimer's disease. *Cell. Mol. Neurobiol.* **31**, 939–948 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors are grateful for support from the following sources: the Alzheimer's Association (grant 4079 to G.S.B.); the Owens Family Foundation (G.S.B.); the Cure Alzheimer's Fund (G.S.B., C.G.G.); NIH/NIGMS training grant T32 GM008136, which funded part of J.M.N.'s PhD training; NIH/NIA grant R01 AG033069 (C.G.G.); and the German Federal Department of Science and Technology grant 03IS2211F (H.-U.D.). We also thank H. Dawson and M. Vitek of Duke University for providing the tau-knockout mice. This work fulfilled part of the requirements for the PhD earned by J.M.N. at the University of Virginia. The technical assistance of A. Spano, H.-H. Ludwig, E. Scheel and K. Schulz is gratefully acknowledged.

**Author Contributions** J.M.N. performed most of the biochemical and cell biological experiments; S.S. was the principal force behind the experiments involving hAPP<sub>SL</sub>/hQC and TBA2.1/tau-knockout mice, and was aided by B.H.-P. and H.C.; A.S. and T.W. fractionated and analysed human brain extracts; E.S., K.T. and B.W. performed the peri-hippocampal injection experiments; A.H. and C.G.G. produced and characterized the M64 and M87 antibodies; R.R. and K.R. performed the electrophysiology experiments; A.A., W.J. and S.G. performed and analysed the immunohistochemical experiments on TBA2.1 and tau-knockout/TBA2.1 mice; G.S.B. and H.-U.D. initiated and directed the project; G.S.B. was the principal writer of the paper; all of the authors participated in the design and analysis of experiments, and in editing of the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to G.S.B. ([gsb4g@virginia.edu](mailto:gsb4g@virginia.edu)) or H.-U.D. (Hans-Ulrich.Demuth@probiodrug.de).



# Clonally related visual cortical neurons show similar stimulus feature selectivity

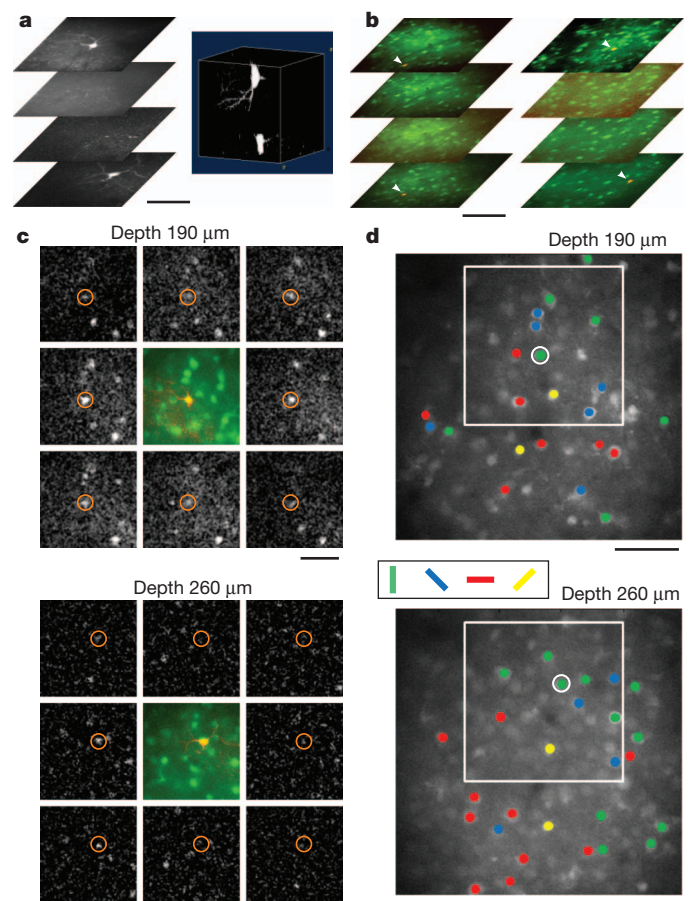
Ye Li<sup>1,2\*</sup>, Hui Lu<sup>1,2\*</sup>, Pei-lin Cheng<sup>1</sup>, Shaoyu Ge<sup>3</sup>, Huatai Xu<sup>4</sup>, Song-Hai Shi<sup>4</sup> & Yang Dan<sup>1,2</sup>

A fundamental feature of the mammalian neocortex is its columnar organization<sup>1</sup>. In the visual cortex, functional columns consisting of neurons with similar orientation preferences have been characterized extensively<sup>2–4</sup>, but how these columns are constructed during development remains unclear<sup>5</sup>. The radial unit hypothesis<sup>6</sup> posits that the ontogenetic columns formed by clonally related neurons migrating along the same radial glial fibre during corticogenesis<sup>7</sup> provide the basis for functional columns in adult neocortex<sup>1</sup>. However, a direct correspondence between the ontogenetic and functional columns has not been demonstrated<sup>8</sup>. Here we show that, despite the lack of a discernible orientation map in mouse visual cortex<sup>4,9,10</sup>, sister neurons in the same radial clone exhibit similar orientation preferences. Using a retroviral vector encoding green fluorescent protein to label radial clones of excitatory neurons, and *in vivo* two-photon calcium imaging to measure neuronal response properties, we found that sister neurons preferred similar orientations whereas nearby non-sister neurons showed no such relationship. Interestingly, disruption of gap junction coupling by viral expression of a dominant-negative mutant of Cx26 (also known as Gjb2) or by daily administration of a gap junction blocker, carbenoxolone, during the first postnatal week greatly diminished the functional similarity between sister neurons, suggesting that the maturation of ontogenetic into functional columns requires inter-cellular communication through gap junctions. Together with the recent finding of preferential excitatory connections among sister neurons<sup>11</sup>, our results support the radial unit hypothesis and unify the ontogenetic and functional columns in the visual cortex.

To identify clonally related sister cells, we used a green fluorescent protein (GFP)-expressing retrovirus, previously shown to label isolated ontogenetic columns of excitatory neurons<sup>7,11,12</sup>. The retrovirus was injected into the right ventricle *in utero* at embryonic day 15–17 (E15–17; see Methods), the beginning of neurogenesis in cortical layer 2/3 (ref. 13). At postnatal day 12–17 (P12–17, soon after eye opening), *in vivo* two-photon imaging<sup>14,15</sup> was performed in the primary visual cortex (V1) of injected mice under anaesthesia. A low density of GFP-labelled neurons was observed in layer 2/3 ( $1.1 \pm 0.9$  (standard deviation) per animal, within an imaging window  $\sim 500 \mu\text{m}$  in diameter at cortical depths up to  $400 \mu\text{m}$ ,  $n = 181$  neurons, 161 mice). In some cases ( $n = 52$ ), we found a pair of GFP-labelled neurons aligned nearly vertically (Fig. 1a, b), with no other GFP neurons nearby, suggesting that they were clonally related sister cells. Although large tangential dispersion has been observed in some clonally related cells<sup>16</sup>, here we focused on GFP-labelled neuronal pairs with  $<120 \mu\text{m}$  horizontal separation (see Methods; Supplementary Fig. 1).

To examine the functional properties of layer 2/3 neurons, we injected the calcium indicator dye Oregon Green BAPTA-1 AM (OGB-1) into a region encompassing the GFP-labelled cell pair. Orientation and direction selectivity of OGB-1-loaded neurons was measured with drifting grating stimuli (100% contrast, spatial frequency 0.02–0.03

cycles per degree, temporal frequency, 1–2 Hz) presented through the contralateral eye. The mapping was made at two cortical depths to include both GFP-labelled neurons in the sister pair. We found that 875 of the 2,286 OGB-1-loaded neurons (38%) showed significant increases in intracellular calcium in response to the grating stimuli



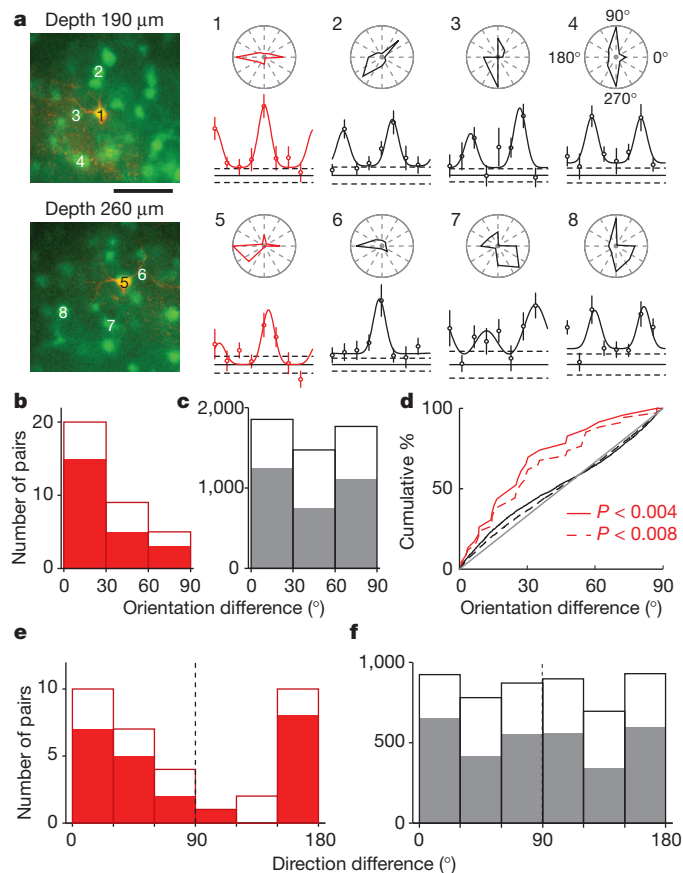
**Figure 1 | Two-photon imaging of clonally related sister cells and nearby layer 2/3 neurons.** **a**, Left, fluorescence images at 190–260  $\mu\text{m}$  from pia, showing two GFP-expressing cells that were nearly vertically aligned. Scale bar, 50  $\mu\text{m}$ . Right, three-dimensional reconstruction of the GFP pair. **b**, Two more examples of GFP cell pairs (arrowheads). Red, GFP; green, OGB-1. Left, 250–310  $\mu\text{m}$  from pia; right, 150–240  $\mu\text{m}$ . Scale bar, 100  $\mu\text{m}$ . **c**, Single condition maps of fluorescence change ( $\Delta F$ ), computed by averaging the images during each stimulus and subtracting baseline (grey screen) for the experiment in **a**. Central panels, GFP (red) and OGB-1 (green) labelling. Red circles, GFP-labelled cells. Scale bar, 50  $\mu\text{m}$ . **d**, Orientation maps with visually responsive cells coloured according to their preferred orientations for a larger imaging area. White box, region shown in **c**. Scale bar, 50  $\mu\text{m}$ .

<sup>1</sup>Division of Neurobiology, Department of Molecular and Cell Biology, Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720, USA. <sup>2</sup>Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. <sup>3</sup>Department of Neurobiology & Behavior, State University of New York at Stony Brook, Stony Brook, New York 11794, USA. <sup>4</sup>Developmental Biology Program, Memorial Sloan-Kettering Cancer Centre, 1275 York Avenue, New York, New York 10065, USA.

\*These authors contributed equally to this work.

(see Methods; Fig. 1c). Among these visually driven neurons, 75% (657/875) showed significant orientation selectivity ( $P < 0.05$ , Hotelling's T-squared test), comparable to previous studies in rodent visual cortex<sup>4,17</sup>. Notably, nearby neurons often preferred different orientations with no apparent spatial organization (Fig. 1c, d), consistent with previous findings of a 'salt-and-pepper' arrangement of orientation preferences in rodent visual cortex<sup>4,10</sup>.

Comparing the response properties of the sister cells, however, we found that they often preferred similar orientations (Fig. 1c, d, circles). To quantify this relationship, we fitted the tuning curve of each visually driven neuron with a double Gaussian function (Fig. 2a and Supplementary Fig. 2) to identify its preferred orientation ( $\theta$ ). The functional similarity between each cell pair was measured by the difference between their preferred orientations ( $\Delta\theta$ , varying between  $0^\circ$  and  $90^\circ$ ). Of the 34 sister pairs in which both neurons were visually driven, 20 pairs (59%) preferred similar orientations ( $0^\circ < \Delta\theta < 30^\circ$ ), and only 5 pairs (15%) preferred near orthogonal orientations ( $60^\circ < \Delta\theta < 90^\circ$ ; Fig. 2b). The distribution of  $\Delta\theta$  was significantly non-uniform ( $P = 0.0071$ , Kolmogorov–Smirnov test), with a strong bias towards 0. In contrast, for pairs of non-sister neurons with horizontal distance  $< 120 \mu\text{m}$ , the distribution of  $\Delta\theta$  was largely flat (Fig. 2c), significantly



**Figure 2 | Orientation and direction preference of sister neurons.**

**a**, Orientation tuning of the GFP pair (red) and several nearby cells (indicated by numbers) in the experiment shown in Fig. 1a, c, d. Circle, mean; vertical line,  $\pm$  s.e.m.; curve, fitted double Gaussian function. Horizontal lines, baseline (solid)  $\pm$  s.e.m. (dashed). Inset above, polar plot of orientation tuning. Scale bar,  $50 \mu\text{m}$ . **b**, Histogram distribution of difference in preferred orientation ( $\Delta\theta$ ) between sister cells, for all visually driven (open bars) and orientation selective (filled bars) pairs. **c**, Histogram of  $\Delta\theta$  between non-sister (GFP/non-GFP and non-GFP/non-GFP) pairs, for all visually driven (open) and orientation-selective (filled) pairs. **d**, Cumulative distribution of  $\Delta\theta$  for sister and non-sister pairs. Red, sisters ( $P$  values, difference from uniform distribution); black, non-sisters. Dashed, all visually driven pairs; solid, orientation-selective pairs. Grey, diagonal line. **e**, **f**, Histograms of difference in preferred direction between sister cells (**e**) and non-sister cells (**f**).

different from the sister pairs (Fig. 2d;  $P = 0.018$ ). The slight biases of the distribution for non-sister pairs towards both  $0^\circ$  and  $90^\circ$  were caused by the overrepresentation of cardinal orientations (horizontal and vertical) in mouse V1 soon after eye opening<sup>17</sup>.

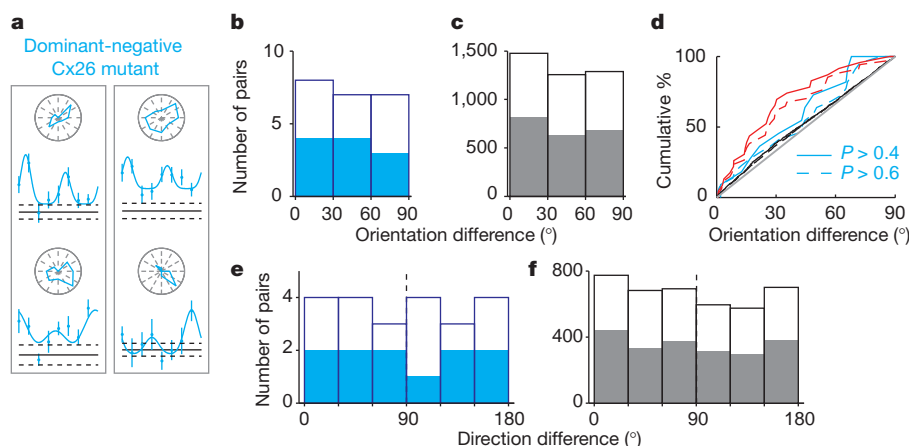
When we restricted the analysis to cell pairs with significant orientation selectivity ( $n = 23$ ), the distribution of  $\Delta\theta$  for sister pairs showed a more marked bias towards 0 (Fig. 2b, filled bars), significantly different from both the uniform distribution ( $P = 0.0038$ ) and the distribution for orientation-selective non-sister pairs (Fig. 2c, filled bars;  $P = 0.034$ ). Furthermore, even with the sparse labelling, our population of sister pairs could still be contaminated by GFP-labelled neurons from separate but neighbouring radial clones. Thus, the similarity between true sisters could be even stronger than that observed here.

In addition to similar orientation tuning, the sister neurons also showed a modest tendency to prefer similar directions. When their difference in preferred direction was plotted over the range  $0$ – $180^\circ$ , we found more pairs falling between  $0^\circ$  and  $90^\circ$  (21/34 visually driven, 14/23 orientation-selective pairs) than between  $90^\circ$  and  $180^\circ$  (Fig. 2e), although the difference was not statistically significant ( $P = 0.11$  for visually driven,  $P = 0.26$  for significantly tuned pairs, bootstrap). For the non-sister pairs, the distribution was largely symmetrical (Fig. 2f).

We next explored the mechanism that confers sister neurons with similar functional properties. Previous studies in developing cortical slices have revealed spontaneous co-activation of neurons within discrete, radially oriented domains spanning multiple cortical layers, which is mediated by gap junction coupling between the neurons<sup>18</sup>. These domains are comparable to the radial clones in shape and size, and they could provide a blueprint for the functional columns by influencing the formation and fine tuning of chemical synapses. To test this idea, we examined the effect of disrupting gap junction coupling between cortical neurons on orientation tuning. Among all the genes encoding the gap junction protein connexin, Cx26 was found to be highly expressed in developing neocortex and strongly associated with interneuronal coupling<sup>19</sup>. We thus injected *in utero* a retrovirus expressing a mutant Cx26, with a threonine in the third transmembrane helix (T135) replaced by alanine<sup>20</sup> (Cx26(T135A)-T2A-EGFP). Because the mutant Cx26 forms closed gap junction channels and exerts a trans-dominant-negative effect on other connexins<sup>21</sup>, it provides a useful tool for selective disruption of gap junction communication in a small number of cortical neurons (which in this case were labelled with enhanced (E)GFP). When we measured orientation tuning at P12–17, we found that 35% (585/1,674) of the OGB-1-loaded neurons were visually responsive, among which 72% (423/585) showed significant orientation selectivity ( $P < 0.05$ , Hotelling's T-squared test), similar to the mice injected with retrovirus expressing GFP alone. Thus, expression of the mutant Cx26 in a small number of neurons caused no global disruption of V1 responses. However, in contrast to the GFP control (Fig. 2b, d), the distribution of  $\Delta\theta$  for sister pairs showed little bias towards 0 (Fig. 3b, d), not significantly different from uniform or the distribution for non-sister pairs (Fig. 3c;  $P > 0.6$ , Kolmogorov–Smirnov test). We also found no clear tendency of the sister cells to prefer similar directions (Fig. 3e, f). Thus, in addition to their roles in prenatal neuronal proliferation and migration<sup>22</sup>, gap junctions may also be required for coordinating postnatal functional development of sister cells, through either electrical coupling or intercellular exchanges of small molecules<sup>23</sup>.

In addition to the selective disruption of gap junction coupling in a small number of neurons with retrovirus, we also tested the effect of systemic application of a gap junction blocker, carbenoxolone (CBX), in the mice injected with retrovirus expressing GFP alone. As gap junction coupling between cortical neurons declines rapidly in the second postnatal week<sup>24</sup>, we injected CBX (intraperitoneally,  $10$ – $20 \text{ mg kg}^{-1}$ ) daily for the first postnatal week to disrupt interneuronal communication during early postnatal development but not at the time of imaging. We found that CBX injection similarly disrupted the functional





**Figure 3 | Effect of expressing dominant-negative mutant of Cx26 in sister neurons.** **a**, Two example experiments showing tuning curves of sister pairs (each pair in a box) in mice injected with retrovirus containing Cx26(T135A)-T2A-EGFP. **b**, Histogram of difference in preferred orientation ( $\Delta\theta$ ) between sister neurons for all visually driven (open) and orientation-selective (filled) pairs in mice expressing mutant Cx26. **c**, Histogram of  $\Delta\theta$  between non-sisters

in these mice. **d**, Cumulative distribution of  $\Delta\theta$  for sister pairs expressing mutant Cx26 (cyan;  $P$  values, comparison with uniform distribution), those expressing GFP only (red, same as Fig. 2d), and for non-sister pairs in mice injected with retrovirus containing Cx26(T135A)-T2A-EGFP (black). Dashed, all visually driven pairs; solid, orientation-selective pairs. **e**, **f**, Histograms of difference in preferred direction between sisters (**e**) and non-sisters (**f**).

similarity between sister neurons (Supplementary Fig. 3). However, CBX injection also caused an overall reduction in the percentage of visually driven neurons (225/1,121, 20%). This suggests that gap junctions may have an important role in early postnatal development of normal visual responses, although with systemic application of CBX it is difficult to rule out its potential non-specific effects on cortical development.

Intracortical excitatory connections are highly non-random<sup>25</sup>, organizing the neurons into fine-scale subnetworks<sup>26</sup>. A recent study showed that sister neurons in the same radial clone are much more connected to each other than to nearby non-sister neurons<sup>11</sup>, suggesting that the radial clones may provide a basis for subnetwork organization. The high connectivity between sister cells should also contribute to their functional similarity as observed in our study. However, inputs from the sister cells alone are likely to be insufficient to determine stimulus selectivity, as each neuron receives inputs from  $\sim 1,000$  other neurons, whereas each radial clone only consists of tens of neurons<sup>6</sup>. Other factors, such as common inputs to the sister cells, may also have important roles. In mouse V1, layer 2/3 neurons with similar orientation tuning are shown to be preferentially interconnected<sup>10</sup>. A significant fraction of these neurons may be sister cells, exhibiting similar orientation tuning (Fig. 2) and preferential connectivity<sup>11</sup>.

Although the columnar structure has long been thought to be a fundamental organizational principle of the neocortex, the existence of a basic processing unit has remained controversial<sup>8</sup>. Although the anatomical minicolumns observed in adult cortex<sup>27</sup> are believed to arise from ontogenetic columns, the relationship between the functional columns and mini/ontogenetic columns remained speculative<sup>1</sup>. Our results demonstrate a direct correspondence between them in V1, at least in superficial layers where neurons are most orientation selective<sup>28</sup>. Contrary to the notion of random organization, our study shows that orientation tuning is organized in columns even in rodent visual cortex. The fine spatial scale of ontogenetic columns may also explain the extraordinary precision of the orientation map in cat visual cortex<sup>4,29</sup>. The interspecies difference in macroscopic cortical organization may be due to differences in the horizontal connections between ontogenetic columns<sup>1</sup>, which can lead to either a smoothly varying map or apparent salt-and-pepper organization<sup>30</sup>. Thus, our results support the view that the ontogenetic columns, rather than the macroscopic functional columns, constitute the basic units of cortical processing.

## METHODS SUMMARY

Retrovirus was injected into the right ventricle of each mouse embryo *in utero* at E15–17. At P12–17, the injected mice were anaesthetized with urethane (1 g kg<sup>-1</sup>)

and chlorprothixene (5 mg kg<sup>-1</sup>), sometimes supplemented with isoflurane (0.5–1% in O<sub>2</sub>). A 1.5-mm-diameter craniotomy was made above V1 for two-photon imaging. After the GFP-expressing cells were identified, nearby layer 2/3 neurons were labelled with OGB-1 via bolus loading. For measuring orientation tuning, 8–12 repeats of drifting sinusoidal gratings were presented in 8 directions spanning 0–360° in a pseudo-random sequence. Neuronal responses were measured at the two depths where the GFP-expressing neurons were found. To quantify orientation and direction preference, we fitted each measured tuning curve by a double Gaussian function. To block gap junctions, 10–20 mg kg<sup>-1</sup> CBX was injected daily (intraperitoneally) over the first postnatal week.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 13 October 2011; accepted 5 April 2012.

Published online 2 May 2012.

- Mountcastle, V. B. The columnar organization of the neocortex. *Brain* **120**, 701–722 (1997).
- Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160**, 106–154 (1962).
- Bonhoeffer, T. & Grinvald, A. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* **353**, 429–431 (1991).
- Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P. & Reid, R. C. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* **433**, 597–603 (2005).
- White, L. E. & Fitzpatrick, D. Vision and cortical map development. *Neuron* **56**, 327–338 (2007).
- Rakic, P. Specification of cerebral cortical areas. *Science* **241**, 170–176 (1988).
- Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714–720 (2001).
- Horton, J. C. & Adams, D. L. The cortical column: a structure without a function. *Phil. Trans. R. Soc. Lond. B* **360**, 837–862 (2005).
- Schuetz, S., Bonhoeffer, T. & Hubener, M. Mapping retinotopic structure in mouse visual cortex with optical imaging. *J. Neurosci.* **22**, 6549–6559 (2002).
- Ko, H. et al. Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
- Yu, Y. C., Bultje, R. S., Wang, X. & Shi, S. H. Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature* **458**, 501–504 (2009).
- Cepko, C. L. et al. Studies of cortical development using retrovirus vectors. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 265–278 (1990).
- Polleux, F., Dehay, C. & Kennedy, H. The timetable of laminar neurogenesis contributes to the specification of cortical areas in mouse isocortex. *J. Comp. Neurol.* **385**, 95–116 (1997).
- Denk, W., Strickler, J. H. & Webb, W. W. 2-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
- Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl Acad. Sci. USA* **100**, 7319–7324 (2003).
- Walsh, C. & Cepko, C. L. Clonal dispersion in proliferative layers of developing cerebral cortex. *Nature* **362**, 632–635 (1993).
- Rochefort, N. L. et al. Development of direction selectivity in mouse cortical neurons. *Neuron* **71**, 425–432 (2011).
- Yuste, R., Peinado, A. & Katz, L. C. Neuronal domains in developing neocortex. *Science* **257**, 665–669 (1992).



19. Nadarajah, B., Jones, A. M., Evans, W. H. & Parnavelas, J. G. Differential expression of connexins during neocortical development and neuronal circuit formation. *J. Neurosci.* **17**, 3096–3111 (1997).
20. Beahm, D. L. *et al.* Mutation of a conserved threonine in the third transmembrane helix of  $\alpha$ - and  $\beta$ -connexins creates a dominant-negative closed gap junction channel. *J. Biol. Chem.* **281**, 7994–8009 (2006).
21. Rouan, F. *et al.* Trans-dominant inhibition of connexin-43 by mutant connexin-26: implications for dominant connexin disorders affecting epidermal differentiation. *J. Cell Sci.* **114**, 2105–2113 (2001).
22. Elias, L. A. & Kriegstein, A. R. Gap junctions: multifaceted regulators of embryonic cortical development. *Trends Neurosci.* **31**, 243–250 (2008).
23. Kandler, K. & Katz, L. C. Coordination of neuronal activity in developing visual cortex by gap junction-mediated biochemical communication. *J. Neurosci.* **18**, 1419–1427 (1998).
24. Peinado, A., Yuste, R. & Katz, L. C. Extensive dye coupling between rat neocortical neurons during the period of circuit formation. *Neuron* **10**, 103–114 (1993).
25. Song, S., Sjostrom, P. J., Reigl, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, e68 (2005).
26. Yoshimura, Y., Dantzker, J. L. & Callaway, E. M. Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868–873 (2005).
27. Rockel, A. J., Hiorns, R. W. & Powell, T. P. The basic uniformity in structure of the neocortex. *Brain* **103**, 221–244 (1980).
28. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
29. Ohki, K. *et al.* Highly ordered arrangement of single neurons in orientation pinwheels. *Nature* **442**, 925–928 (2006).
30. Koulakov, A. A. & Chklovskii, D. B. Orientation preference patterns in mammalian visual cortex: a wire length minimization approach. *Neuron* **29**, 519–527 (2001).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by National Institutes of Health (NIH) grant R01 EY018861 and NSF grant 22250400-42533 (to Y.D.), and NIH grants R01 DA024681 and R21NS072483 (to S.-H.S.). We thank L. E. White and S. D. Van Hooser for comments on the manuscript, A. Kwan and S. D. Van Hooser for help with two-photon imaging techniques and analysis and Y. Gu for help in making retrovirus.

**Author Contributions** Y.L. performed the two-photon imaging experiments and data analysis. H.L., Y.L. and P.-I.C. performed *in utero* virus injection. S.G., H.X. and S.-H.S. provided the viral vectors. Y.L., H.L. and Y.D. designed the experiments and wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to Y.D. ([ydan@berkeley.edu](mailto:ydan@berkeley.edu)).

## METHODS

**Retroviral infection.** Replication-incompetent retrovirus expressing either GFP alone or a loss-of-function mutant of connexin 26 and EGFP (Cx26(T135A)-T2A-EGFP)<sup>20</sup> was produced as previously described<sup>31</sup>. Uterine horns of E15–17 gestation stage pregnant female C57BL/6 mice (Charles River Laboratories) were exposed in a clean environment. Retrovirus (0.5–1  $\mu$ l) with fast green (2.5 mg ml<sup>-1</sup>; Sigma) was injected into the right embryonic cerebral ventricle at a speed of 150 nl s<sup>-1</sup>, controlled by a microinjection pump (WPI). After injection, the peritoneal cavity was washed with warm saline solution containing antibiotics, the uterine horns were replaced, and the wound was closed. Both male and female mice were used in the experiments. All experimental procedures were approved by the Animal Care and Use Committee at the University of California, Berkeley.

**Two-photon imaging.** Mice were anaesthetized by intraperitoneal injection of urethane (1 g kg<sup>-1</sup>) and chlorprothixene (5 mg kg<sup>-1</sup>), in some cases supplemented with isoflurane (0.5–1% in O<sub>2</sub>). Body temperature was maintained at 37 °C using a feedback heating pad. The head was secured using a stainless steel head plate affixed onto the skull using super glue and dental cement. A 1.5-mm-diameter craniotomy was performed at the location of the primary visual cortex (0–1 mm anterior to the lambda suture, 2–2.5 mm lateral of the midline). The dura was left intact. The cortical surface was constantly irrigated with an extracellular solution (in mM: 135 NaCl, 5 KCl, 5 HEPES, 1.8 CaCl<sub>2</sub> and 1 MgCl<sub>2</sub>, at pH 7.3 and ~285 mOsm).

The neocortical neurons were labelled with calcium indicator dye via bolus loading<sup>32</sup>. The dye solution consists of 1 mM OGB-1, 10% dimethylsulphoxide (DMSO), 2% (wt/vol) Pluronic F-127 in HEPES-buffered saline (in mM: 150 NaCl, 2.5 KCl and 10 HEPES). Multiple injections of the dye solution were made in adjacent regions at a depth of ~200  $\mu$ m. The experiment began 1 h after the dye injection. The two-photon microscope (Movable Objective Microscope, Sutter Instrument) was controlled using the ScanImage software<sup>33</sup>. The intensity of the excitation from a tunable femtosecond laser (Wideband, Tsunami Mode-Locked Ti: Sapphire Laser, Spectra-Physics) was controlled by a Pockels cell (350-80-LA-02; Conoptics). The excitation was focused using a 40 $\times$ /0.8 NA infrared objective (LUMPLFLN, Olympus). Fluorescence was collected after a dichroic mirror (670DCXXR, Chroma) using a pentagon-style detector into green and red channels, with respective emission filters (FF01-510/84-25, Semrock; HQ610/75, Chroma) and photomultiplier tubes (GaAsP H10770PA-40 and multi-alkali R6357, Hamamatsu). Different excitation wavelengths were used to measure fluorescence of OGB-1 (800 nm) and GFP (900 nm). To image OGB-1 during visual stimulation, frames of 512  $\times$  512 pixels were acquired continuously every 1.5–1.8 s.

**Identification of sister neurons.** As shown in Supplementary Fig. 1, the distribution of the horizontal distance between each pair of GFP-expressing neurons showed a prominent peak at <100  $\mu$ m and a long tail. We chose a relatively conservative criterion of 120  $\mu$ m (dashed line, Supplementary Fig. 1) for the pair of neurons to be considered sisters. In a previous study, the mean horizontal spread of the newly born neurons at P14–16 was found to be >500  $\mu$ m (ref. 11). Thus, it is possible that some GFP-expressing sister neurons with large horizontal separation were excluded from our analysis. However, because increasing the criterion distance is likely to increase the probability of false positives (misclassification of non-sister pairs as sister pairs), in this study we chose to focus on cell pairs with small horizontal separations.

Of course, even with this relatively conservative criterion, one cannot exclude the possibility that some GFP-expressing neurons arising from different radial clones were close to each other simply by chance and were thus misclassified as sisters. As the distribution of  $\Delta\theta$  for non-sisters was largely flat (Fig. 2c, f), the contamination of the sister-pair population by non-sister pairs should cause broadening of the observed distribution. Thus the similarity between true sisters in orientation and direction preference may be even stronger than that shown in Fig. 2b, e.

**Visual stimulation.** Visual stimuli were generated with a PC computer containing a NVIDIA GeForce 6600 graphics board and presented with a XENARC 700V LCD monitor (19.7 cm  $\times$  12.1 cm, 960  $\times$  600 pixels, 75 Hz refresh rate, 300 cd m<sup>-2</sup> maximum luminance, gamma corrected with custom software) located 14 cm from the left eye, positioned such that the receptive fields of the imaged neurons were at the centre of the monitor. For measuring orientation tuning and direction selectivity of V1 neurons, full-field drifting gratings (100% contrast, 1–2 Hz, 0.02–0.03 cycles per degree) were presented in 8 directions (separated by 45°) in a pseudorandom sequence. Each stimulus was 5 s in duration with a 5 s interstimulus interval. After each block of 8 drifting gratings, 5 s of blank stimulus (grey screen) was presented to measure the baseline activity. A total of 8–12 blocks were presented in each experiment.

**Data analysis.** Images were analysed with custom software written in Matlab. Small horizontal drift over time was corrected by measuring correlation at different pixel offsets and realigning the images according to the best match. Cells were identified by the experimenter. For each frame, the fluorescence value of each cell was computed by averaging all pixels in a circle (radius, 12 pixels, 5.5  $\mu$ m) centred on the soma.

The response to each stimulus was calculated as  $\frac{\Delta F}{F} = \frac{(F_{\text{STIM}}) - F_0}{F_0}$ , where  $F_{\text{STIM}}$  is the average response across all frames when the stimulus is on, and  $F_0$  is the average response during the final 3 s of the interstimulus period before stimulus onset. Cells were considered visually responsive if the responses to visual stimuli were different from the response to blank by ANOVA test (at  $P < 0.05$ ). Among cells that were visually responsive, cells with significant orientation selectivity were identified by plotting the response in each trial as a point in orientation space:  $S(\theta_i)e^{\frac{2\pi i \theta_i}{180}}$ , where  $S(\theta_i)$  is the raw  $\frac{\Delta F}{F}$  at direction  $\theta_i$ . Cells were considered orientation selective if the mean of the cloud of points was significantly different from (0,0) by Hotelling's T-squared test (at  $P < 0.05$ ).

To identify the preferred orientation and direction of each cell, the responses to drifting gratings were fit with a 2-peak Gaussian function:

$$R(\theta) = R_{\text{OFFSET}} + R_{\text{PREF}} e^{-\frac{\text{ang}(\theta - \theta_{\text{PREF}})^2}{2\sigma^2}} + R_{\text{OPP}} e^{-\frac{\text{ang}(\theta + 180 - \theta_{\text{PREF}})^2}{2\sigma^2}}$$

where  $R_{\text{OFFSET}}$  is a constant offset,  $\theta_{\text{PREF}}$  is the preferred direction,  $R_{\text{PREF}}$  is the above-offset response to the preferred direction,  $R_{\text{OPP}}$  is the above-offset response to the opposite direction,  $\sigma$  is the tuning width and  $\text{ang}(x) = \min(x, x - 360, x + 360)$ , which wraps angular difference values onto the interval 10° to 180°.

- van Praag, H. *et al.* Functional neurogenesis in the adult hippocampus. *Nature* **415**, 1030–1034 (2002).
- Garaschuk, O., Milos, R. I. & Konnerth, A. Targeted bulk-loading of fluorescent indicators for two-photon brain imaging *in vivo*. *Nature Protocols* **1**, 380–386 (2006).
- Pologruto, T., Sabatini, B. & Svoboda, K. ScanImage: Flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003).

# A global synthesis reveals biodiversity loss as a major driver of ecosystem change

David U. Hooper<sup>1</sup>, E. Carol Adair<sup>2,3</sup>, Bradley J. Cardinale<sup>4</sup>, Jarrett E. K. Byrnes<sup>2</sup>, Bruce A. Hungate<sup>5</sup>, Kristin L. Matulich<sup>6</sup>, Andrew Gonzalez<sup>7</sup>, J. Emmett Duffy<sup>8</sup>, Lars Gamfeldt<sup>9</sup> & Mary I. O'Connor<sup>2,10</sup>

Evidence is mounting that extinctions are altering key processes important to the productivity and sustainability of Earth's ecosystems<sup>1–4</sup>. Further species loss will accelerate change in ecosystem processes<sup>5–8</sup>, but it is unclear how these effects compare to the direct effects of other forms of environmental change that are both driving diversity loss and altering ecosystem function. Here we use a suite of meta-analyses of published data to show that the effects of species loss on productivity and decomposition—two processes important in all ecosystems—are of comparable magnitude to the effects of many other global environmental changes. In experiments, intermediate levels of species loss (21–40%) reduced plant production by 5–10%, comparable to previously documented effects of ultraviolet radiation and climate warming. Higher levels of extinction (41–60%) had effects rivaling those of ozone, acidification, elevated CO<sub>2</sub> and nutrient pollution. At intermediate levels, species loss generally had equal or greater effects on decomposition than did elevated CO<sub>2</sub> and nitrogen addition. The identity of species lost also had a large effect on changes in productivity and decomposition, generating a wide range of plausible outcomes for extinction. Despite the need for more studies on interactive effects of diversity loss and environmental changes, our analyses clearly show that the ecosystem consequences of local species loss are as quantitatively significant as the direct effects of several global change stressors that have mobilized major international concern and remediation efforts<sup>9</sup>.

A variety of global changes are driving rates of species extinction that greatly outpace background rates in the fossil record<sup>10,11</sup>. If these trends continue, projections suggest that within 240 years Earth may face the sixth mass extinction<sup>12</sup>. Such projections have prompted hundreds of experiments examining how different components of biodiversity affect ecosystem processes that sustain the provisioning of goods and services to society. Syntheses of these experiments have made it clear that plant biodiversity loss will reduce plant production and alter decomposition<sup>5,6</sup>. However, it is uncertain how the sizes of these effects compare with the direct effects of other types of environmental change, such as changing atmospheric composition, climate warming and nutrient pollution, that also threaten ecosystem functioning<sup>13–15</sup>. This uncertainty has generated wide-ranging speculation about how strongly biodiversity loss might affect humanity<sup>16,17</sup>.

Here we report the results of a large data synthesis in which we compared the effects of species loss against other drivers of environmental change. We focus on primary production and decomposition because these major biological processes influence carbon storage and other ecosystem services, and illustrate the breadth of sensitivity of ecosystem processes to changes in species richness<sup>2,6,18</sup>. We took two approaches in our analyses. First, we statistically summarized existing meta-analyses that have estimated the mean effect size of experimental

manipulations of a variety of environmental changes on primary production (biomass production by plants) and decomposition (mass loss of plant litter) in a variety of ecosystems around the world (Tables 1 and 2). We compared these environmental effect sizes to the estimated effects of species loss derived from a database we constructed using 192 peer-reviewed publications on experiments that manipulated species richness and examined the effects on ecosystem processes (see Methods). This approach allows comparison among a wide range of environmental changes, but has the limitation that it evaluates the effects of environmental and diversity changes measured by different researchers using different organisms and ecosystems. To complement our summary of meta-analyses, we also summarized the results of 16 experiments that simultaneously manipulated plant species richness in factorial combination with some other environmental change (elevated CO<sub>2</sub>, nutrient pollution, etc). Although a far smaller data set, analysis of factorial experiments allowed two additional comparisons: (1) effect sizes of diversity loss versus other environmental changes, within experiments focusing on identical ecosystems; and (2) effect sizes of diversity loss under current versus projected environmental conditions. We assessed a breadth of projections of local species loss because estimates vary widely for magnitudes of global species extinctions (Supplementary Table 1). Similarly, species losses at local scales most relevant to biodiversity and ecosystem functioning (BEF) experiments and ecosystem services (m<sup>2</sup> to watersheds) probably do not bear a one-to-one relationship with global extinctions (complete loss of a species from the planet) and may respond nonlinearly to multiple environmental changes<sup>10,19,20</sup>.

Our analyses suggest that biodiversity loss in the 21st century could rank among the major drivers of ecosystem change. Experiments to date have shown that effects of plant species richness on biomass production are nonlinear and saturating (Fig. 1). Our analysis suggests that in areas where local species loss this century falls within the lower range of projections (1–20%), negligible effects on biomass production will result, and changes in species richness will rank low relative to the effects projected for other environmental changes (Fig. 1 and Table 1). Where actual losses fall within intermediate projections (21–40%), however, species loss is expected to reduce biomass production by 5–10% below the most diverse mixtures (based on exponentiation of log response ratios (LRR):  $e^{-0.05} = 0.951$ ,  $e^{-0.107} = 0.898$ ). This effect is comparable in magnitude to the effects of ultraviolet radiation and climate warming on plant production (Fig. 1 and Table 1).

Where losses fall within higher projections of extinction (41–60%), the effects of species loss rank with those of many other drivers of environmental change, such as warming, ozone and acidification (Fig. 1). The mid-point of this range, fifty per cent species loss, is a benchmark at the upper end of 21st century projections of global extinctions, but is a common estimate at the local scale in heavily-affected

<sup>1</sup>Department of Biology, Western Washington University, Bellingham, Washington 98225-9160, USA. <sup>2</sup>National Center for Ecological Analysis and Synthesis, 735 State Street, Suite 300, Santa Barbara, California 93101, USA. <sup>3</sup>Rubenstein School of Environment and Natural Resources, Aiken Center, University of Vermont, Burlington, Vermont 05405, USA. <sup>4</sup>School of Natural Resources & Environment, University of Michigan, Ann Arbor, Michigan 48109-1041, USA. <sup>5</sup>Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011, USA. <sup>6</sup>Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697-2525, USA. <sup>7</sup>Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montréal, Québec H3A 1B1, Canada. <sup>8</sup>Virginia Institute of Marine Science, College of William and Mary, Gloucester Point, Virginia 23062, USA. <sup>9</sup>Department of Biological and Environmental Sciences, University of Gothenburg, Box 461, SE-405 30 Göteborg, Sweden. <sup>10</sup>Department of Zoology, University of British Columbia, 2370-6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada.



**Table 1 | Effects of species richness and environmental changes on primary productivity for the broad meta-analysis and factorial diversity crossed with environment experiments.**

Factor	Broad meta-analysis				Factorial experiments			
	$N_S, N_{obs}$	LRR	LCI	UCI	$N_S, N_{obs}$	LRR	LCI	UCI
Primary producer diversity								
50% loss	60, 145	−0.144	−0.175	−0.112	10, 15	−0.168	0.104	−0.439
Avg. mono.	73, 299	−0.332	−0.378	−0.285	16, 30	−0.458	−0.259	−0.658
Best mono.	62, 241	0.159	0.116	0.203	13, 29	−0.136	0.067	−0.338
Other factors								
Acidif.	1, 12	−0.186	−0.342	−0.020				
+Ca	1, 31	0.351	−0.105	0.820	1, 1	0.256	−0.781	1.293
+CO <sub>2</sub>	6, 3076	<b>0.217</b>	<b>0.207</b>	<b>0.227</b>	3, 5	0.070	−0.400	0.539
Drought	1, 20	−0.616	−0.892	−0.342	3, 5	0.215	−0.255	0.686
+N	6, 2895	<b>0.310</b>	<b>0.192</b>	<b>0.428</b>	+N-med 3, 13	0.155	−0.155	0.466
					+N-high 2, 21	0.434	0.165	0.703
+N +CO <sub>2</sub>	1, 252	<b>0.694</b>	<b>0.622</b>	<b>0.766</b>				
+N +P	1, 941	<b>0.964</b>	<b>0.894</b>	<b>1.034</b>	5, 8	0.586	0.215	0.958
+Ozone	4, 2162	−0.149	−0.161	−0.137				
+P	2, 766	<b>0.239</b>	<b>0.175</b>	<b>0.302</b>	1, 1	1.216	0.177	2.254
Plant inv.	1, 144	<b>0.514</b>	<b>0.447</b>	<b>0.581</b>				
+Ultraviolet	2, 432	−0.082	−0.107	−0.057				
Warming	1, 1064	0.116	0.078	0.154	1, 1	−0.529	−1.568	0.510

$N_S$ , total number of studies—references listed in Supplementary Table 2, except for diversity effects for the broad meta-analysis which come from the database on biodiversity and ecosystem functioning (BEF) experiments<sup>6</sup>;  $N_{obs}$ , total number of observations across all meta-analyses or experiments; LRR, log response ratio; LCI, lower 95% confidence interval; UCI, upper 95% confidence interval. Bold values indicate bootstrapped mean LRRs and confidence intervals in the broad meta-analysis (see Supplementary Fig. 2). Treatment factors: 50% loss; Avg. mono., average monoculture; Best mono., best monoculture (see Methods for calculation of LRRs). In all of these, negative values indicate that species loss causes a decline in productivity rates. Acidif., acidification; Ca, calcium; CO<sub>2</sub>, carbon dioxide; N, nitrogen (for terrestrial N addition, 'low' rates were  $\leq 3 \text{ g m}^{-2}$  (no such factorial experiments), medium (med.) were  $> 3$  and  $\leq 15 \text{ g m}^{-2}$ , and high were  $> 15 \text{ g m}^{-2}$ ); P, phosphorus; Plant inv., plant invasion.

landscapes that have experienced  $>90\%$  habitat loss<sup>21</sup>. A 50% species loss is expected to reduce biomass production by an average of 13% ( $e^{-0.144}$ ; Table 1), an effect consistent across terrestrial, freshwater and marine ecosystems (Supplementary Fig. 1). For comparison, elevated CO<sub>2</sub> experiments have produced greater overall magnitudes of changes in biomass (+24%). This average, however, combines studies performed in diverse natural systems as well as in agricultural monocultures. Experiments performed in multi-species communities have shown the effects of CO<sub>2</sub> on production of +12–13% (ref. <sup>22</sup>, Supplementary Fig. 2, Tot<sub>multi</sub> under Elevated CO<sub>2</sub>)—on par with projected effects of 50% species loss. Similarly, the average effect of nitrogen (N) on plant biomass production depended on N addition rates. Rates of N addition similar to intensive agricultural fertilization had effects on production (+54% for factorial experiments) that were greater than those of intermediate or high levels of species loss. However, the magnitude of effects of high species loss on production was comparable to those of intermediate (+17%; Table 1) or low (rare in terrestrial experiments, Supplementary Fig. 3)<sup>23</sup> rates of N addition. Thus, the magnitude of the effects of high species loss on production also seems to be comparable to those of increased nitrogen deposition, a well-recognized environmental problem<sup>24,25</sup>.

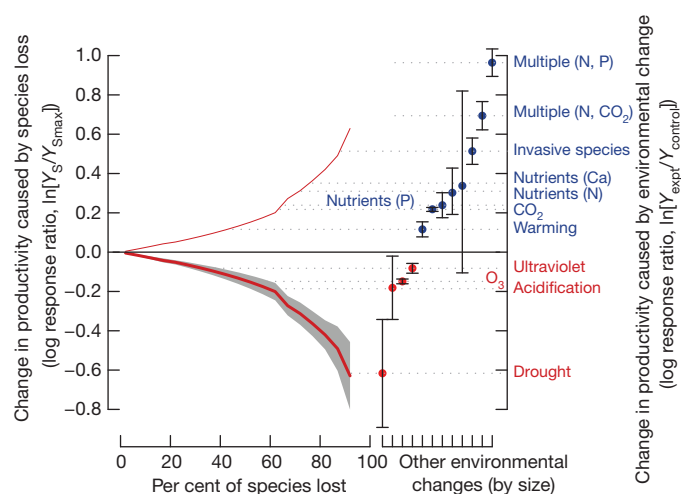
To rival the environmental changes that have had the greatest documented effect on primary production (for example, heavy nutrient pollution, invasive species, drought), species loss would need to exceed that of prior mass extinctions ( $\geq 75\%$  species loss). This scenario is

unlikely to be realized globally in the coming century, but could occur for some types of organisms (for example, vertebrates) within 240–540 years if current rates of extinction continue<sup>12</sup>. It may also occur for a variety of organisms at local scales where human activities heavily affect land use. If such a scenario were realized for plants, biomass production in natural systems would be expected to decline by an average of one third, exceeding the effects of all other environmental changes except invasive species, drought and interactions among multiple pollutants (N, P, CO<sub>2</sub>) applied in combination (Fig. 1). However, uncertainty around the effects of extinction grows large as the fraction of species loss increases, in part because the identity and biological traits of surviving species have an increasingly large effect on biomass production. The importance of species identity is most apparent from examining extreme cases where experiments reduce diverse communities to single species. Whereas the average effect of reducing diversity to a monoculture is a 28% loss in production, the distribution ranged from −68% to +62% of production compared to the most diverse mixture. Similarly, if one can conserve the most productive monocultures, these outperformed the most diverse mixtures by an average 17%, but ranged from −40% to +132% (Table 1; Supplementary Discussion, Productivity section). These values span the range of plausible effects for nearly all other environmental changes, and underscore the large variation in possible outcomes of extinction that can result from functional differences among species and ecosystems<sup>4</sup>.

**Table 2 | Effects of species richness and environmental changes on decomposition from the broad meta-analysis.**

Factor	$N_S, N_{obs}$	LRR	LCI	UCI	Response variable
Litter diversity					
50% loss	24, 39	0.023	−0.062	0.108	Mixed
Avg. mono.	31, 67	0.034	−0.130	0.199	Mixed
Best mono.	21, 39	0.266	0.153	0.378	Mixed
Consumer diversity					
50% loss	22, 52	−0.074	−0.155	0.008	Mixed
Avg. mono.	22, 55	−0.235	−0.359	−0.111	Mixed
Best mono.	19, 49	−0.056	−0.190	0.077	Mixed
Other factors					
+CO <sub>2</sub>	1, 101	−0.020	−0.041	0.010	Mass loss
Eutrophication	1, 6	0.250	−0.180	0.660	Microbial breakdown rate
Plant inv.	2, 62	<b>0.729</b>	<b>0.677</b>	<b>0.782</b>	Decomposition rate, microbial breakdown rate
Acidification	1, 5	−0.830	−1.600	−0.520	Microbial breakdown rate
+N	2, 520	−0.023	−0.046	0.000	Decomposition rate, mass loss

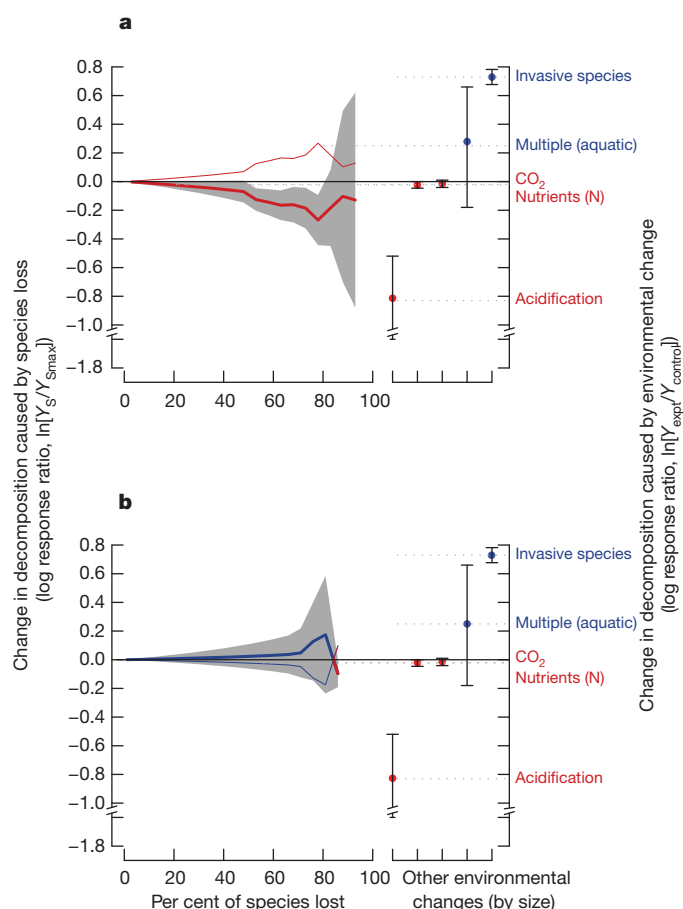
Abbreviations as in Table 1. Response variable, biomass types and response variables used to calculate LRR value. Mixed: microbial respiration rate, decomposition rate, mass loss rate, feeding rate. Bold values indicate bootstrapped mean LRRs and confidence intervals.



**Figure 1 | Changes in primary production as a function of per cent local species loss.** Effects of species loss on primary production from 62 studies (379 observations). Thick red line, lower productivity as species richness decreases; grey bands and black error bars, 95% confidence intervals. The thin red line shows the inverse of the thick red line to allow comparison of effect magnitudes with environmental changes with positive effects. Dotted grey lines show the mean effect of each environmental change for comparison with the effect of richness. Right axis, effects of other environmental changes. Blue is for increases and red for decreases in productivity (Table 1 and Supplementary Table 2).

Analysis of experiments that manipulated species richness in factorial combination with other environmental changes generally reinforced our conclusions from the broader meta-analysis (Table 1; Supplementary Discussion, Productivity section). Within intermediate projections of species loss (21–40%), the effects of species loss on plant biomass production equalled or exceeded the effects of elevated CO<sub>2</sub>, and rivalled the effect of drought (Table 1 and Supplementary Fig. 4). Interactions between species loss and environmental changes are important for understanding net effects on ecosystem processes, because both will often occur simultaneously (environmental changes rank among the major drivers of species loss<sup>13,26,27</sup>). We compared the effects of species loss (average monoculture metric) under experimental conditions with the effects of species loss under control environmental conditions to investigate potential interactions between these drivers of ecosystem processes (Supplementary Figs 5 and 6). The available evidence indicates that diversity effects were independent of many environmental changes (interactions were not detectably different from zero). The exception was N addition, which led to smaller average effects of diversity under elevated than control conditions ( $P = 0.043$ , when weighted by  $n$ ; Supplementary Fig. 5). However, the scarcity of studies meant we found three or fewer experiments for any change other than fertilization. Clearly, this is a critical topic for future research.

Both environmental changes and species loss had smaller effects on decomposition than on production. However, the effects of consumer species loss on decomposition were comparable to the effects of some major forms of environmental change. Loss of litter consumer richness reduced decomposition rates by ~8% for mid-ranges of projected extinction, giving rise to effects that were comparable in magnitude to elevated CO<sub>2</sub> (–2%) and nitrogen pollution (–2%), although smaller than the effects of multiple nutrient addition in aquatic systems, acidification and plant invasion (Fig. 2 and Table 2). The effects of consumer loss were more pronounced and consistent in freshwater, where the majority of experiments have taken place<sup>6</sup>, than in terrestrial ecosystems (–12% versus –7%, respectively, for a 50% loss scenario, Supplementary Fig. 1). In contrast to the effects of consumer species loss, loss of litter diversity did not alter average rates of decomposition (Fig. 2 and Table 2; Supplementary Discussion, Decomposition section). Because species loss reduced primary productivity more than



**Figure 2 | Changes in decomposition as a function of per cent local species loss.** **a**, Effects of detrital consumer diversity on decomposition from 19 studies (54 observations). **b**, Effects of plant litter diversity on decomposition from 22 studies (60 observations). Thick red lines, slower decomposition rates as species richness decreases; thick blue lines, higher decomposition rate as species richness decreases; grey bands and black error bars, 95% confidence intervals. Thin coloured lines, dotted grey lines, axes and colour coding as in Fig. 1. See also Table 2 and Supplementary Table 3.

decomposition, future species loss could limit the capacity for carbon uptake and storage in the biosphere<sup>18</sup>.

In summary, we have shown that species loss ranks among the major drivers of primary production and decomposition—key processes involved in the carbon cycle and the provisioning of many ecosystem services<sup>11,18</sup>. Refining these estimates for key ecosystem services will require a better understanding of how realistic extinction scenarios interact with other forms of environmental change in influencing multiple ecosystem processes<sup>4,16,26</sup>. Even so, the range of effects caused by species loss spanned the range of plausible outcomes for nearly all other drivers of environmental change. And the average effects of local extinction were comparable in magnitude to numerous other global change stressors that have already mobilized major international concern and remediation efforts. As such, our study provides a quantitative basis for integrating consequences of species loss into assessments to be conducted by the Intergovernmental Science Policy Platform for Biodiversity and Ecosystem Services<sup>28</sup>.

## METHODS SUMMARY

To quantify how species loss affects primary production and decomposition, we used the database of ref. 6 that summarized 192 studies (574 experiments) through 2009 that manipulated species richness and measured the effects on ecosystem processes. We extracted experiments describing (1) how species richness of primary producers influenced producer biomass and (2) how primary producer or consumer richness affected decomposition of litter. We then calculated two log response

ratios (LRRs) for each experiment:  $\ln(Y_{\text{Ave mono}}/Y_{\text{Smax}})$  and  $\ln(Y_{\text{Best mono}}/Y_{\text{Smax}})$  where  $Y_{\text{Smax}}$  was production or decomposition in the most diverse mixture in an experiment,  $Y_{\text{Ave mono}}$  is the average value of the monocultures, and  $Y_{\text{Best mono}}$  is the value in the most productive or fastest decomposing monoculture (for considerations, see Methods and Supplementary Discussion<sup>29</sup>). When possible, we also fit data from each study to a power function  $\ln(Y_{\text{S}}/Y_{\text{Smax}}) = a + b \times \ln(S)$ . Parameter estimates were used to produce the nonlinear species loss curves in Figs 1, 2 and Supplementary Fig. 4, and the 50% loss scenario in Tables 1 and 2.

To gather data on how other forms of environmental change affect production and decomposition, we searched the ISI Web of Science for published meta-analyses (see Methods). From each paper or publicly available data set<sup>23</sup>, we extracted response ratios ( $RR = Y_{\text{exp}}/Y_{\text{control}}$ , where  $Y_{\text{exp}}$  is the response variable in the experimental treatment), number of studies, and estimates of variance (Data Thief III, Version 1.5). We calculated the overall mean LRR and 95% confidence interval for each treatment via bootstrapping using skew normal distributions<sup>30</sup>.

We also identified 16 factorial experiments to directly compare productivity LRRs for diversity to other forms of environmental change in the same experiment. Where manipulations either reduced or enhanced resources, we changed the sign of the LRRs to allow comparison to the broader meta-analysis. We analysed LRRs using mixed models in SYSTAT v.12, with environmental change as a fixed effect and study as a random effect.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 9 January; accepted 13 April 2012.**

**Published online 2 May 2012.**

- Loreau, M., Naeem, S. & Inchausti, P. *Biodiversity and Ecosystem Functioning: Synthesis and perspectives* (Oxford Univ. Press, 2002).
- Hooper, D. U. *et al.* Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol. Monogr.* **75**, 3–35 (2005).
- Tilman, D. Ecological consequences of biodiversity: a search for general principles. *Ecology* **80**, 1455–1474 (1999).
- Wardle, D. A., Bardgett, R. D., Callaway, R. M. & Van der Putten, W. H. Terrestrial ecosystem responses to species gains and losses. *Science* **332**, 1273–1277 (2011).
- Balvanera, P. *et al.* Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* **9**, 1146–1156 (2006).
- Cardinale, B. J. *et al.* The functional role of producer diversity in ecosystems. *Am. J. Bot.* **98**, 572–592 (2011).
- Stachowicz, J. J., Bruno, J. F. & Duffy, J. E. Understanding the effects of marine biodiversity on communities and ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **38**, 739–766 (2007).
- Perrings, C. *et al.* Ecosystem services, targets, and indicators for the conservation and sustainable use of biodiversity. *Front. Ecol. Environ.* **9**, 512–520 (2011).
- IPCC. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Core Writing Team, Pachauri, R. K. & Reisinger, A.) (IPCC, 2007).
- Sala, O. E. *et al.* Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1774 (2000).
- Millennium Ecosystem Assessment. *Ecosystems and Human Well-being: Biodiversity Synthesis* (World Resources Institute, 2005).
- Barnosky, A. D. *et al.* Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).
- Chapin, F. S. III *et al.* Consequences of changing biodiversity. *Nature* **405**, 234–242 (2000).
- Grace, J. B. *et al.* Does species diversity limit productivity in natural grassland communities? *Ecol. Lett.* **10**, 680–689 (2007).
- Paquette, A. & Messier, C. The effect of biodiversity on tree productivity: from temperate to boreal forests. *Glob. Ecol. Biogeogr.* **20**, 170–180 (2011).
- Srivastava, D. S. & Vellend, M. Biodiversity-ecosystem function research: is it relevant to conservation? *Annu. Rev. Ecol. Evol. Syst.* **36**, 267–294 (2005).
- Rockström, J. *et al.* A safe operating space for humanity. *Nature* **461**, 472–475 (2009).
- Díaz, S., Wardle, D. A. & Hector, A. in *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective* (eds Naeem, S. *et al.*) Ch. 11 149–166 (Oxford Univ. Press, 2009).
- Pereira, H. M. *et al.* Scenarios for global biodiversity in the 21st century. *Science* **330**, 1496–1501 (2010).
- Brook, B. W., Sodhi, N. S. & Bradshaw, C. J. A. Synergies among extinction drivers under global change. *Trends Ecol. Evol.* **23**, 453–460 (2008).
- Ewers, R. M. & Didham, R. K. Confounding factors in the detection of species responses to habitat fragmentation. *Biol. Rev. Camb. Philos. Soc.* **81**, 117–142 (2006).
- Wang, X. Effects of species richness and elevated carbon dioxide on biomass accumulation: a synthesis using meta-analysis. *Oecologia* **152**, 595–605 (2007).
- Elser, J. J. *et al.* Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* **10**, 1135–1142 (2007).
- Vitousek, P. M. *et al.* Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* **7**, 737–750 (1997).
- Carpenter, S. R. *et al.* Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **8**, 559–568 (1998).
- Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11**, 1351–1363 (2008).
- Suding, K. N. *et al.* Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Glob. Change Biol.* **14**, 1125–1140 (2008).
- Larigauderie, A. & Mooney, H. A. The Intergovernmental science-policy Platform on Biodiversity and Ecosystem Services: moving a step closer to an IPCC-like mechanism for biodiversity. *Curr. Opin. Environ. Sust.* **2**, 9–14 (2010).
- Schmid, B., Hector, A., Saha, P. & Loreau, M. Biodiversity effects and transgressive overyielding. *J. Plant Ecol.* **1**, 95–102 (2008).
- Johnson, N. J. Modified *t* tests and confidence intervals for asymmetrical populations. *J. Am. Stat. Assoc.* **73**, 536–544 (1978).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the authors of the original data sets, both published and unpublished, who shared data with us. We thank the National Center for Ecological Analysis and Synthesis (NCEAS) for funding the workshop “Biodiversity and the Functioning of Ecosystems: Translating Results from Model Experiments into Functional Reality”. Support for NCEAS comes from University of California Santa Barbara and the National Science Foundation. J.E.K.B., E.C.A. and M.I.O. had NCEAS post-doctoral fellowships, J.E.D. had support from NSF OCE-1031061; B.J.C. had support from NSF DEB-1046121; L.G. was supported by grant 621-2009-5457 from the Swedish Research Council VR; A.G. is supported by the Canada Research Chair Program and NSERC.

**Author Contributions** All authors contributed to the design of the study, data interpretation and manuscript editing; B.J.C. and K.L.M. developed the database of biodiversity and ecosystem functioning experiments; D.U.H., E.C.A., J.E.K.B., B.J.C. and K.L.M. collected additional data and performed statistical analyses. E.C.A., J.E.K.B., B.J.C., B.A.H. and D.U.H. drafted the figures and D.U.H. wrote the initial draft.

**Author Information** The biodiversity and ecosystem functioning database is deposited with the National Center for Ecological Analysis and Synthesis (<http://kn.b.eoinformatics.org/knb/metacat/nceas.984/nceas>). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to D.U.H. ([hooper@biol.wvu.edu](mailto:hooper@biol.wvu.edu)).



## METHODS

**Biodiversity and ecosystem functioning database.** To quantify the effects of species loss on biomass production and decomposition, we used the data set of ref. 6. This data set summarizes 192 peer-reviewed papers published through 2009 reporting results from 574 experiments that manipulated species richness and measured the effects on various ecosystem processes. We extracted the subset of experiments that examined (1) how species richness of primary producers influenced producer biomass accumulation and (2) how richness of producer litter, or richness of litter consumers, affected decomposition rates (Tables 1 and 2). For multi-year studies, we used only data from the last year as this was least likely to be influenced by transient responses. For each experiment, we calculated two log response ratios (LRRs):  $\ln(Y_{\text{Ave mono}}/Y_{\text{Smax}})$  and  $\ln(Y_{\text{Best mono}}/Y_{\text{Smax}})$  where  $Y_{\text{Smax}}$  was production or decomposition in the most diverse mixture in an experiment,  $Y_{\text{Ave mono}}$  is the average value of the monocultures, and  $Y_{\text{Best mono}}$  is the value in the most productive or fastest decomposing monoculture. Both LRRs quantify the net effect of species loss going from the most to least diverse communities, but differ in their assumptions about the sequence of extinction. For productivity only, best monoculture values were restricted to communities where increasing plant diversity increased production (that is, where the average monoculture LRR was positive)<sup>6</sup>. Statistical issues may bias the effects of best monocultures<sup>29</sup>; because this topic is controversial, however, we use the best monoculture metric primarily to illustrate the range of potential process responses, particularly in heavily managed ecosystems.

Log ratios like those described above are frequently used to summarize diversity effect sizes<sup>6</sup>, in part because they can be calculated for most experiments. However, these metrics represent extreme scenarios of local diversity loss that are not likely to be realized in many natural communities. Therefore, we also ran a more comprehensive analysis on the subset of experiments that included at least three levels of species richness. For these experiments, we fit the mean value of the response at each level of richness  $S$  to the power function:  $\ln(Y_S/Y_{\text{Smax}}) = a + b \times \ln(S)$ . Prior meta-analyses lend much stronger support to saturating models of diversity effects (for example, Michaelis–Menten or power functions) compared to linear or exponential fits<sup>6</sup>. We used the power function here because it gave a good fit and provided a balance between simplicity and generality (mean  $R^2 = 0.71$  for productivity and 0.30 for decomposition, compared to mean  $R^2 = 0.73$  and 0.29 for Michaelis–Menten). After obtaining parameter estimates for each experiment, we calculated the effect of species loss on production and decomposition across all levels of per cent loss that we could interpolate within an individual experiment. We calculated the log response ratio  $\ln(Y_S/Y_{\text{Smax}})$  at 5% increments of species loss, where  $Y_S$  is the value at  $S$  species ( $< S_{\text{max}}$ ). The distribution of log ratios was estimated by bootstrapping, and means and 95% confidence intervals were plotted in Fig. 1. The 50% loss scenario in Tables 1 and 2 came from these estimates.

**Meta-analysis comparison.** To quantify how other forms of environmental change have an effect on production and decomposition, we collated data published in past syntheses and data analyses. These studies typically manipulated abiotic conditions consistent with accepted scenarios of environmental change for the factors at hand (for example, doubling of atmospheric  $\text{CO}_2$ , ref. 9; Supplementary Fig. 2).

Productivity: we searched ISI Web of Science for meta-analyses that examined the effects of global change factors on biomass production (search terms: [product\* AND meta-analysis] OR [biomass AND meta-analysis]). Because more recent meta-analyses often have extensive reference overlap with earlier meta-analyses of the same environmental effect, we only used meta-analyses published after 2005 to maximize independence across studies. We found 18 meta-analyses summarizing 67 LRRs showing how various aspects of environmental change influence primary production in marine, freshwater and terrestrial ecosystems (Supplementary Table 2). LRRs were calculated as  $\ln(\text{mean treatment biomass}/\text{mean control biomass})$ . From each meta-analysis, we extracted the LRR, number of observations, and the associated error measurement from text, tables or digitized figures (Data Thief III, Version 1.5), or calculated them directly where data were freely available<sup>23</sup>.

We found LRR values for 12 forms of environmental change: acidification, calcium (Ca) additions, elevated  $\text{CO}_2$ , drought, plant invasion, nitrogen (N) additions, phosphorus (P) additions, N + P additions, N additions + elevated  $\text{CO}_2$ , elevated ozone, elevated ultraviolet radiation, and warming (Table 1 and Supplementary Table 2). If a treatment was represented by only one LRR value, then the reported LRR and associated confidence intervals were used in our analysis (Table 1). If a treatment was represented by more than one LRR value, we calculated the overall mean LRR and confidence interval for each treatment via bootstrapping from skew normal distributions<sup>30</sup>. Distributions were resampled 10,000 times to generate an overall mean and lower/upper confidence interval using the 'fGarch' package of R version 2.12.2. Bootstrapped means and

confidence intervals are compared with the means and confidence intervals from the original data sources in Supplementary Fig. 2.

Decomposition: to limit reference overlap, we searched for decomposition meta-analyses published after 2000, using the search terms [decomp\* AND meta-analysis] in the ISI Web of Science. We found five meta-analyses, resulting in seven LRRs of a treatment effect on decomposition in freshwater and/or terrestrial ecosystems (Table 2 and Supplementary Table 3). LRR values were available for five different treatments: acidification, elevated  $\text{CO}_2$ , plant invasion, N additions and eutrophication (multiple nutrient additions in aquatic ecosystems). We extracted data and calculated mean LRR and associated confidence intervals as described above for productivity.

**Environment crossed with species richness manipulations.** We complemented our summary of meta-analyses with a more focused analysis that compared the effects of species richness to the effects of other forms of environmental change when both were manipulated simultaneously in the same experiment. To do this, we extracted records from the ref. 6 database for experiments that manipulated species richness and some component of environmental change in factorial combination. We only had sufficient data to assess effects of diversity and environmental manipulations on biomass production (16 studies, Supplementary Table 2): +calcium<sup>31</sup>, + $\text{CO}_2$  (refs 32–34), water availability ('drought')<sup>35–37</sup>, nitrogen addition<sup>31,32,38</sup>, phosphorus addition<sup>39</sup>, multiple nutrient addition<sup>40–44</sup>, and warming<sup>45</sup>. In our statistical analyses, we also included effects of light manipulation<sup>46</sup>, although the explicit link to global environmental change is less clear for this factor, so it is not shown in figures. For each study, we calculated the suite of diversity LRRs previously described, as well as the effect of the environmental manipulations at maximum species richness (Table 1). For experiments where manipulations either reduced or enhanced resources (for example, nutrient or water availability), we changed the sign of the LRRs appropriately so that magnitudes of effects could be compared on a scale similar to environmental changes assessed in the broader meta-analysis. We analysed LRRs using mixed models in SYSTAT v.12 (SYSTAT, Inc.) with environmental change as a fixed effect and study as a random effect. We compared equally weighted results to analyses where we weighted LRRs by sample size  $(n_1 \times n_2)/(n_1 + n_2)$ ; results were qualitatively similar, unless otherwise noted.

- Rixen, C., Huovinen, C., Huovinen, K., Stöckli, V. & Schmid, B. A plant diversity × water chemistry experiment in subalpine grassland. *Perspect. Plant Ecol.* **10**, 51–61 (2008).
- Reich, P. B. *et al.* Plant diversity enhances ecosystem responses to elevated  $\text{CO}_2$  and nitrogen deposition. *Nature* **410**, 809–810 (2001).
- Maestre, F. T. & Reynolds, J. F. Biomass responses to elevated  $\text{CO}_2$ , soil heterogeneity and diversity: an experimental assessment with grassland assemblages. *Oecologia* **151**, 512–520 (2007).
- Stocker, R., Körner, C., Schmid, B., Niklaus, P. A. & Leadley, P. W. A field study of the effects of elevated  $\text{CO}_2$  and plant species diversity on ecosystem-level gas exchange in a planted calcareous grassland. *Glob. Change Biol.* **5**, 95–105 (1999).
- Mulder, C. P. H., Uliassi, D. D. & Doak, D. F. Physical stress and diversity–productivity relationships: the role of positive species interactions. *Proc. Natl Acad. Sci. USA* **98**, 6704–6708 (2001).
- Rixen, C. & Mulder, C. P. H. Improved water retention links high species richness with increased productivity in arctic tundra moss communities. *Oecologia* **146**, 287–299 (2005).
- Weninger, E. J. & Inouye, R. S. Insect community response to plant diversity and productivity in a sagebrush–steppe ecosystem. *J. Arid Environ.* **72**, 24–33 (2008).
- Wacker, L., Baudois, O., Eichenberger-Glinz, S. & Schmid, B. Diversity effects in early- and mid-successional species pools along a nitrogen gradient. *Ecology* **90**, 637–648 (2009).
- Striebel, M., Behl, S. & Stibor, H. The coupling of biodiversity and productivity in phytoplankton communities: consequences for biomass stoichiometry. *Ecology* **90**, 2025–2031 (2009).
- Fridley, J. D. Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia* **132**, 271–277 (2002).
- Lanta, V. & Lepš, J. Effect of functional group richness and species richness in manipulated productivity–diversity studies: a glasshouse pot experiment. *Acta Oecol.* **29**, 85–96 (2006).
- Smith, A. & Allcock, P. J. The influence of species diversity on sward yield and quality. *J. Appl. Ecol.* **22**, 185–198 (1985).
- Boyer, K. E., Kertész, J. S. & Bruno, J. F. Biodiversity effects on productivity and stability of marine macroalgal communities: the role of environmental context. *Oikos* **118**, 1062–1072 (2009).
- von Felten, S. & Schmid, B. Complementarity among species in horizontal versus vertical rooting space. *J. Plant Ecol.* **1**, 33–41 (2008).
- De Boeck, H. J. *et al.* Biomass production in experimental grasslands of different species richness during three years of climate warming. *Biogeosciences* **5**, 585–594 (2008).
- Fridley, J. D. Diversity effects on production in different light and fertility environments: an experiment with communities of annual plants. *J. Ecol.* **91**, 396–406 (2003).

# CORRECTIONS & AMENDMENTS

---

## ERRATUM

doi:10.1038/nature11138

### **Erratum: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability**

Ben B. B. Booth, Nick J. Dunstone, Paul R. Halloran,  
Timothy Andrews & Nicolas Bellouin

*Nature* **484**, 228–232 (2012)

The accepted date should read 08 February 2012, rather than 2011.  
This error has been corrected in the PDF and HTML versions online.

## BLACK HOLES

# Star ripped to shreds

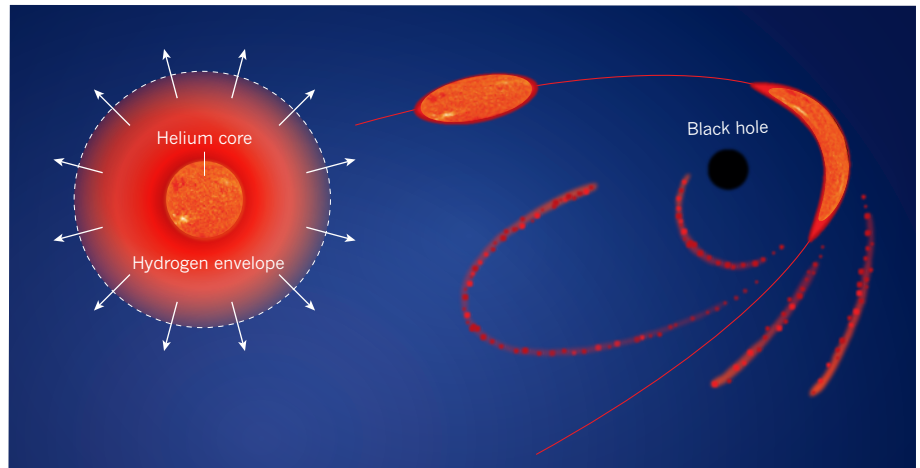
When a star wanders too close to a giant black hole, it can be pulled apart by the black hole's tidal force. One such event offers insight into the properties of both the black hole and the star.

GIUSEPPE LODATO

Astronomers have strong evidence that supermassive black holes, with masses between a million and a billion times that of the Sun, reside in the centre of most galaxies. The evidence comes from observations of the copious amount of radiation that is emitted when these objects pull gas from their immediate vicinity. However, if a black hole's close environment is poor in gas, gas accretion proceeds at a low rate and is not accompanied by significant emission of radiation. Probing such 'dormant' black holes is therefore difficult — unless a tidal-disruption event occurs. Such an event takes place when a star comes close enough to the black hole to be ripped apart by its tidal force. The ensuing stellar debris is accreted by the black hole and produces a characteristic flare. In a paper published on *Nature's* website today, Gezari *et al.*<sup>1</sup> describe how detailed observations of a tidal-disruption event have allowed them to determine the properties of not only the black hole, but also the disrupted star.

Tidal-disruption events are rare. They are expected to occur once every 10,000 years per galaxy. To find them, it is therefore necessary to use large astronomical surveys in which thousands of galaxies are regularly observed. Gezari and colleagues' discovery comes from one such survey, which is expected to detect roughly one event every two years<sup>2</sup>.

The main significance of this study probably lies in the accurate determination of the properties of the tidal-disruption event, which the authors based on a well-sampled ultraviolet–optical light curve for the flare (a plot that shows the evolution of the flare's brightness over time) and spectroscopic measurements of the system. They find that the supermassive black hole is hosted by a galaxy at redshift 0.17 (corresponding to a distance of approximately 2 billion light years from Earth), and has a mass of about 3 million solar masses. Moreover, on the basis of the shape of the light curve<sup>3</sup> and the absence of hydrogen lines in the spectra of the stellar debris, Gezari *et al.* conclude that the disrupted star was the helium-rich core of a red giant whose hydrogen outer shell had been previously stripped off, possibly by the same tidal force that eventually led to its complete disruption (Fig. 1).



**Figure 1 | Tidal disruption of a star.** The helium-rich core of a red-giant star that had previously lost its hydrogen envelope moves on an almost parabolic orbit (red) towards a supermassive black hole. The sequence of blobs illustrates the progressive distortion of the star's core due to the tidal pull of the black hole. After the point of closest approach to the black hole, the core is completely disrupted, with part of the resulting debris being expelled from the system and part being launched into highly eccentric orbits, eventually falling onto the black hole. Accretion of this debris gives rise to the intense ultraviolet–optical flare that has been observed by Gezari and colleagues<sup>1</sup>.

Gezari and colleagues' observations also imply that the orbit of the star around the black hole was exceptionally tight, with the point of closest approach being only six times the black hole's Schwarzschild radius, which, for a non-rotating black hole, corresponds to its event horizon — the boundary beyond which nothing, not even light, can escape. After the point of closest approach, part of the debris was expelled from the system and part was launched into highly eccentric orbits, falling onto the black hole after approximately two months from closest approach, and producing the observed ultraviolet–optical flare.

During the past year, two other tidal-disruption events caused by supermassive black holes have been detected<sup>4–7</sup>. In those cases, the emission occurred over a wide range of the electromagnetic spectrum, from X-rays to radio waves, because it was produced by a high-energy jet of particles that happened to point almost exactly in our line of sight; such jets are expected to be associated with tidal-disruption events. By contrast, the ultraviolet–optical radiation of tidal-disruption events is usually associated with thermal emission from an accretion disk of stellar debris that forms

around the black hole. However, somewhat surprisingly, in the present case Gezari and colleagues find that the ultraviolet–optical emission does not seem to be caused by an accretion disk, because its emission would have fallen off with time<sup>8</sup> at a different rate from that observed. This finding will provide food for thought for subsequent investigations of these events.

Finally, the observation that the star's closest approach to the black hole corresponds to a distance of only six times the Schwarzschild radius suggests that effects of the general theory of relativity might need to be invoked to describe the system. Would the shape of the light curve be modified by such effects? Current models, including those used by Gezari *et al.*, are generally based on Newtonian dynamics, and hence do not include the effects of general relativity. As a result, the models cannot distinguish between a rotating and a non-rotating black hole. Tidal-disruption events can probe deeply into the gravitational field of the black hole, and so offer the means to test the effects of general relativity and black-hole rotation. The fact that we are now in a position to characterize the light curve and



spectra of such events so accurately, and that the rate of discovery seems to be in line with expectations, means that we should soon be able to answer these fascinating questions. ■

**Giuseppe Lodato** is in the Physics Department, University of Milan,

*I-20133 Milan, Italy.*  
e-mail: [giuseppe.lodato@unimi.it](mailto:giuseppe.lodato@unimi.it)

1. Gezari, S. *et al.* *Nature* <http://dx.doi.org/10.1038/nature10990> (2012).
2. Gezari, S. *et al.* *Astro2010: The Astronomy and Astrophysics Decadal Survey Science White Pap.* 88; preprint at <http://arxiv.org/abs/0903.1107> (2009).
3. Lodato, G., King, A. R. & Pringle, J. E. *Mon. Not. R. Astron. Soc.* **392**, 332–340 (2009).
4. Bloom, J. S. *et al.* *Science* **333**, 203–206 (2011).
5. Burrows, D. N. *et al.* *Nature* **476**, 421–424 (2011).
6. Zauderer, B. A. *et al.* *Nature* **476**, 425–428 (2011).
7. Cenko, S. B. *et al.* preprint at <http://arxiv.org/abs/1107.5307> (2011).
8. Lodato, G. & Rossi, E. M. *Mon. Not. R. Astron. Soc.* **410**, 359–367 (2011).

## VIROLOGY

## Bird flu in mammals

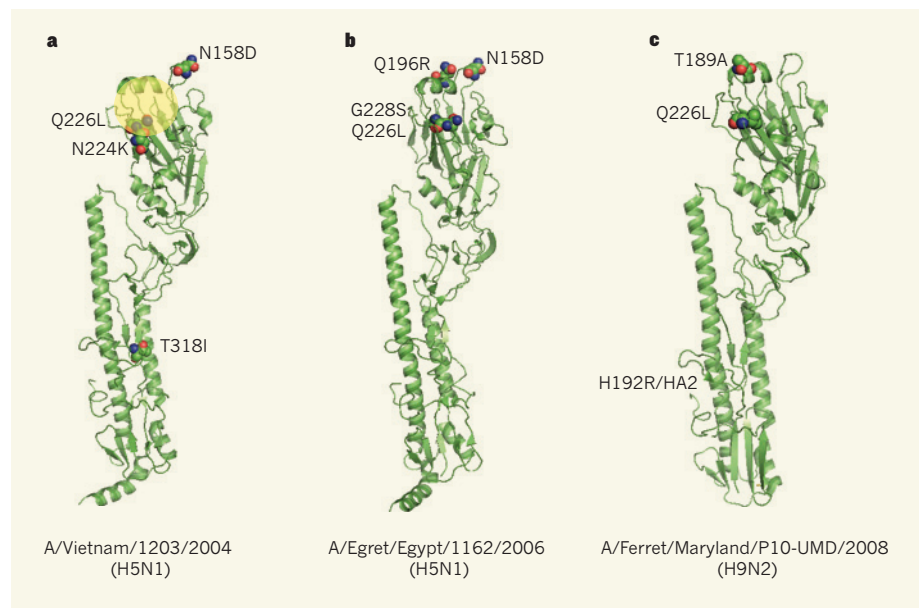
**An engineered influenza virus based on a haemagglutinin protein from H5N1 avian influenza, with just four mutations, can be transmitted between ferrets, emphasizing the potential for a human pandemic to emerge from birds.**

HUI-LING YEN &  
JOSEPH SRIYAL MALIK PEIRIS

Influenza pandemics in humans arise from animal influenza viruses, yet the molecular changes required for an animal virus to be transmitted efficiently between humans are poorly understood. Highly pathogenic H5N1 avian flu viruses have circulated in poultry for more than 16 years, only rarely resulting in human infections. But when people do catch H5N1 bird flu, their disease is of unusual severity, raising concerns that a human H5N1 pandemic might have a catastrophic impact on public health. However, an H5N1 virus that can be efficiently transmitted from human to human has not yet emerged, leading some researchers to question whether these viruses are inherently incapable of acquiring this capacity. In a paper published on *Nature's* website today, Imai *et al.*<sup>1</sup> demonstrate that H5N1 viruses do have the potential to cause a human pandemic. The authors identify mutations in the avian virus that permit viral transmission between ferrets by means of respiratory droplets — the best available model for influenza transmission in humans.

Imai *et al.* focused on the haemagglutinin (HA) protein of a highly pathogenic H5N1 influenza virus, which is involved in binding and fusion of the virus to the cells that it infects; the HA type is used in the viral nomenclature, together with the other influenza surface glycoprotein involved, a neuraminidase (NA). The HA of H5N1 viruses preferentially binds sialic acids in receptors on the surface of avian cells (Siaa2,3), whereas cells of the human upper airway predominantly have another type of sialic acid (Siaa2,6), which is recognized by human influenza viruses. The researchers introduced random mutations into the globular head of the HA molecule, where the receptor-binding domain is located, and searched for mutated viruses that exhibited enhanced binding to Siaa2,6. Using reverse genetics, a technique that allows genetic manipulation of the virus genome, they then made a 'hybrid' H5N1 virus in which the gene encoding one of these mutated H5 HA proteins replaced the HA gene in the H1N1 virus that caused a human pandemic in 2009.

The researchers infected ferrets with this



**Figure 1 | Avian haemagglutinins transmissible in mammals.** The haemagglutinin (HA) protein of influenza determines the type of target cell that the virus can infect. By mutating the site of the HA that binds to sialic acids in receptors on target cells, which differ between avian and mammalian cells, and other protein regions that determine the pH at which virus–cell fusion can occur, researchers have generated viruses that have avian HA proteins that can be transmitted from mammal to mammal. **a**, Imai *et al.*<sup>1</sup> identify four mutations (N158D, N224K, Q226L and T318I) in the HA of a highly pathogenic avian influenza virus, H5N1, that allow a virus with this HA to be transmitted by respiratory droplets between ferrets. The receptor-binding site of the HA is shaded in yellow. **b**, Chen *et al.*<sup>5</sup> also created a virus with an H5N1 HA that can be partially transmitted between ferrets, by introducing three mutations (Q196R, Q226L and G228S) into the H5 HA. The virus used as the basis for this hybrid already contained the N158D mutation that Imai and colleagues also identified in their mutated HA. **c**, Sorrell *et al.*<sup>7</sup> used the HA protein from a low-pathogenic H9N2 avian virus to achieve similar transmissibility. The Q226L mutation was already present in this virus, and two additional mutations (T189A and H192R/HA2) were acquired during ferret infection studies. The H192R/HA2 mutation cannot be shown because the protein is cleaved at a site that lies in front of this amino-acid residue. Full virus names and their subtypes are given below each structure. Mutated amino-acid residues are indicated by red and blue spheres.

hybrid H5N1 virus and, following multiple rounds of infection, isolation of the virus from the upper respiratory tract of the infected animals, and reinfection in animals not previously exposed to the virus, they obtained a virus that can be transmitted efficiently between the animals. Four key amino-acid changes in the HA are associated with ferret respiratory-droplet transmissibility in the authors' virus (Fig. 1a). Three of the mutations (N158D, N224K and Q226L) contribute to Siaa2,6 specificity. The fourth (T318I) lowers the pH at which the protein undergoes a structural change that allows it to release its genetic material into the

cytoplasm of an infected cell by fusion of the viral envelope with intracellular membranes.

There have been many previous attempts to determine whether H5N1 can acquire transmission capacity in ferrets. Two studies<sup>2,3</sup> assessed H5N1 and H3N2 hybrid viruses, and another study<sup>4</sup> introduced mutations into the H5N1 HA that are known to increase Siaa2,6 binding in H2 and H3 haemagglutinin — but neither approach conferred transmissibility by respiratory droplets. One study achieved partial transmissibility by introducing three HA mutations (Q196R, Q226L and G228S) into an H5N1 virus and combining this with the

NA protein of a human seasonal H3N2 virus (Fig. 1b). It is interesting that the two H5 HAs that demonstrate transmissibility among ferrets<sup>1,5</sup> (Fig. 1a,b) contain similar mutations: one at amino-acid residues 158–160, which removes an *N*-linked glycosylation site of the globular head<sup>6</sup>, and the other at residues 221–228, which alters the structure of the loop of the receptor-binding domain<sup>6</sup>.

Another research group previously produced a transmissible H9N2 hybrid virus<sup>7</sup> using the HA and NA of a low-pathogenic avian virus and other genes from a human seasonal H3N2 virus. The HA of this hybrid contains a Q226L mutation, which confers human-like Siaα2,6 binding<sup>7</sup> (Fig. 1c). During the 10 rounds of infection that were needed for this H9N2 virus to acquire transmissibility, it accumulated two additional HA mutations, one located at residue 189 of HA1 (close to the receptor-binding domain) and the other at residue 192 of HA2 (close to the membrane-fusion domain), as well as one NA mutation located in the transmembrane domain. Together, these studies demonstrate that mutations that increase Siaα2,6 binding<sup>1,5,7</sup>, as well as those that stabilize HA structure<sup>1,7</sup>, seem to be functionalities required for the HA of viruses that are transmissible among mammals.

Thus, Imai and colleagues' generation of a respiratory-droplet-transmissible H5N1 virus represents the culmination of a sustained effort by numerous groups to better understand the mechanisms that can confer mammalian transmissibility in avian influenza viruses. However, it is likely that different combinations of HA mutations can achieve the same effects, and further studies are needed to explore this possibility. Furthermore, although transmission of influenza A virus among humans is already known<sup>8,9</sup> to involve the virus HA, NA and basic polymerase protein 2, it is possible that other viral proteins, not explored by Imai and colleagues, also contribute to mammalian transmissibility.

It is intriguing that, although the parent H5N1 virus strain (A/Vietnam/1203/2004) used by Imai *et al.* causes lethal disease in ferrets when they are directly infected, infection with the ferret-transmissible H5N1 virus does not kill the animals. It is possible that the change in receptor binding from Siaα2,3, which is present on alveolar epithelial cells in ferret and human lungs, to Siaα2,6, found on cells of the upper respiratory tract, may change the virus from one that causes alveolar infection, likely to result in more severe disease, to one targeting the upper airways, which is associated with milder symptoms. However, Imai and colleagues did not find that the two viruses differ markedly in their targeting of the upper or lower respiratory regions.

Imai and colleagues' H1N1 virus with an H5 HA is a laboratory creation, but it should not be considered an experimental artefact. Natural emergence of an H5N1–H1N1 hybrid virus is plausible. Some H1N1 and H5N1 viruses readily swap genes with one another *in vitro*<sup>10,11</sup>, generating hybrid viruses. Furthermore, pandemic H1N1 viruses are established in pigs in many parts of the world, and H5N1 viruses have been isolated from pigs<sup>12</sup>, suggesting that opportunities exist for the viruses to combine in these animals.

However, these findings do not only provide further indication that such a virus may arise naturally; they also pave the way for improved influenza surveillance and pandemic preparedness. For example, one of the four mutations reported by Imai *et al.*, N158D, results in a loss of *N*-linked glycosylation, and loss of glycosylation at this residue is increasingly seen among naturally occurring H5N1 isolates. Although the other three mutations have only very rarely been seen in H5N1 isolates from the field, mutations in HA that change the virus from binding Siaα2,3 to binding both Siaα2,3 and Siaα2,6 have been reported in both birds and humans<sup>13</sup>. These findings reinforce the need for focusing even greater attention on

H5N1 infections in humans and other mammals (including pigs). Influenza viruses exist as genetic variants termed quasi-species even within a single clinical specimen, but such genetic diversity may not be fully assessed by conventional sequencing methods. Investigation of mammalian clinical specimens using new deep sequencing methods for these mutations, and for other mutations that confer similar functionality, will allow us to evaluate the extent of H5N1 adaptation that is occurring in mammalian hosts. More broadly, understanding the mutations that confer mammalian transmission of avian influenza viruses will allow better risk assessment of the animal viruses that represent a pandemic threat, and help to select virus strains against which pre-pandemic vaccines should be generated. ■

**Hui-Ling Yen and Joseph Sriyal Malik Peiris** are at the School of Public Health and the Centre of Influenza Research, LKS Faculty of Medicine, The University of Hong Kong, and the HKU-Pasteur Research Centre, Pokfulam, Hong Kong.  
e-mails: malik@hku.hk; hyen@hku.hk

1. Imai, M. *et al.* *Nature* <http://dx.doi.org/10.1038/nature10831> (2012).
2. Jackson, S. *et al.* *J. Virol.* **83**, 8131–8140 (2009).
3. Maines, T. R. *et al.* *Proc. Natl Acad. Sci. USA* **103**, 12121–12126 (2006).
4. Maines, T. R. *et al.* *Virology* **413**, 139–147 (2011).
5. Chen, L.-M. *et al.* *Virology* **422**, 105–113 (2012).
6. Stevens, J. *et al.* *J. Mol. Biol.* **381**, 1382–1934 (2008).
7. Sorrell, E. M., Wan, H., Araya, Y., Song, H. & Perez, D. R. *Proc. Natl Acad. Sci. USA* **106**, 7565–7570 (2009).
8. Belser, J. A., Maines, T. R., Tumpey, T. M. & Katz, J. M. *Expert Rev. Mol. Med.* **12**, e39 (2010).
9. Sorrell, E. *et al.* *Curr. Opin. Virol.* **1**, 635–642 (2011).
10. Cline, T. D. *et al.* *J. Virol.* **85**, 12262–12270 (2011).
11. Octaviani, C. P., Ozawa, M., Yamada, S., Goto, H. & Kawaoka, Y. *J. Virol.* **84**, 10918–10922 (2010).
12. Nidom, C. A. *et al.* *Emerg. Infect. Dis.* **16**, 1515–1523 (2010).
13. Yamada, S. *et al.* *Nature* **444**, 378–382 (2006).